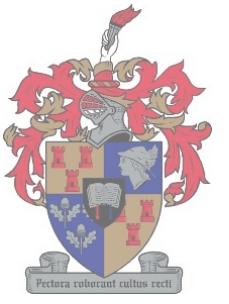
Multi-Bias Decomposition-Based Optimisation for the Extraction of Small-Signal GaAs FET Models

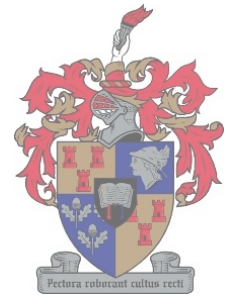
Cornell van Niekerk



Multi-Bias Decomposition-Based Optimisation for the Extraction of Small-Signal GaAs FET Models

Cornell van Niekerk

Dissertation presented for the Degree of Doctor of Philosophy
in Engineering at the University of Stellenbosch



Promoter: Dr. P. Meyer

Date: July 1999

Declaration: *I, the undersigned, hereby declare that the work contained in this dissertation is my own original work, and has not previously in its entirety or in part been submitted at any university for a degree.*

Date: 9 November 1999

Opsomming

Sleutelwoorde: *Parameter Ekstraksie, Optimering, Dekomposisie, MESFET, HEMT, Kleinsein Model*

Die beskikbaarheid van akkurate nie-linêre rekenaar modelle is onmisbaar vir die akkurate en koste effektiewe ontwerp van hoë frekwensie kommunikasie stelsels. Die onttrekking van kleinsein modelle uit gemete s-parameter data is 'n kern stap in die skepping van nie-linêre modelle.

Die tesis beskryf die ontwikkeling van 'n nuwe optimerings algoritme vir die onttrekking van kleinsein modelle uit gemete s-parameters. Die optimerings algoritme is gebasseer op 'n dekomposisie benadering wat die probleem in kleiner subprobleme verdeel en oplos. Die metode is akkuraat, geskik vir hoë dimensionele probleme en kan onakkurate optimerings beginwaardes hanteer. Die sukses van die algoritme is bevestig deur 'n verskeidenheid van toetse op gesimuleerde data, sowel as gemete data van 'n GaAs MESFET and pHEMT. Die resultate dui ook die beperkings van die enkel voorspanning ekstraksies aan. Enkel voorspanning ekstraksies hanteer die data gemeet by verskillende voorspannings as aparte ekstraksie probleme.

Daar word getoon dat die dekomposisie optimeerder uitgebrei kan word om multi-voorspannings probleme te hanteer. In die multi-voorspannings geval word die data gemeet by verskillende voorspannings in een onttrekkings probleem saam gevoeg en die ekstrinsieke elemente van die model word as voorspannings onafhanklik gedefineer. Dit forseer 'n sekere mate van fisiese realiteit op die probleem af. Die multi-voorspannings benadering is robuust en akkuraat en vind ook die waardes van voorheen moeilik bepaalbare model elemente met 'n groot mate van sekerheid. Daar word getoon hoe die ekstraksie onsekerheid afneem soos die hoeveelheid voorspannings in die probleem vermeerder word. Die nuwe algoritme produseer kleinsein modelle wat die s-parameters van die transistor akkuraat modelleer by al die voorspannings.

Abstract

Keywords: *Parameter Extraction, Optimisation, Decomposition, MESFET, HEMT, Small Signal Model*

The availability of accurate nonlinear models is essential for the accurate and cost effective design of high frequency telecommunication systems. A key step in the creation of such models is the extraction of small signal models from measured s-parameter data.

This thesis describes the development and evaluation of a new extraction algorithm. The method makes use of a decomposition-based optimiser which divides the problem into subproblems that are solved separately. The procedure is accurate, suited for handling large problems, and starting value independent. The success of the algorithm is confirmed using a variety of tests. Both simulated and measured data for a GaAs MESFET and pHEMT are used. The test results also show the limitations of single bias extractions. In single bias extractions, the data measured at different bias points are handled as separate extractions.

It is shown that the decomposition-based optimiser can be used to create a new multi-bias extraction algorithm. The multi-bias procedure combines the s-parameters measured at different bias points into one extraction problem, and defines the extrinsic model elements to be bias independent. This enforces a degree of physical reality onto the extraction. Tests prove the algorithm to be robust and accurate. It is able to find previously difficult to determine elements with a high degree of precision. Results are presented that indicate how the extraction uncertainty decreases as the number of bias points used in the multi-bias extraction is increased. The new algorithm produces small signal models that accurately represents the measured s-parameters at all bias points.

Acknowledgements

I would like to express my sincere gratitude to all the people who contributed to the this thesis. In particular I would like to thank my promoter, Dr. Petrie Meyer for his moral and financial support and guidance. Thank you for keeping me on track when I wanted to venture off onto side tracks and keeping me focussed on the problem.

I would also like to thank Dominique Schreurs and Peter Winson, who supplied the measured data with which the final algorithms were tested. I am grateful for your generosity to a stranger and the various discussions that we have since had.

A word of thanks to my friends that contributed to this thesis in various ways. Thanks to Ernst Nolte and Kevin Williams for proof reading this manuscript and to Shelley Sturgeon, whose encouragement during the times that things were not always working, kept me going, even when I wanted to give up.

To my parents a special word of thanks for your understanding, unconditional support and encouragement.

List of Publications

The following publications have resulted from this work:

International Journal Publications:

1. C. van Niekerk, P. Meyer, "A New Approach for the Extraction of an FET Equivalent Circuit from Measured S-Parameters," Microwave and Optical Technology Letters, Vol. 11, No. 5, April 1996, pp. 281-284

2. Van Niekerk, P. Meyer, "Performance and Limitations of Decomposition-Based Parameter-Extraction Procedures for FET Small-Signal Models," IEEE Trans. Microwave Theory and Tech., Vol 46, No. 11, November 1998, pp. 1620-1627

3. C. van Niekerk, P, Meyer, D. Schreurs, P. Winson, "A Robust Integrated Multi-Bias Parameter Extraction Method for MESFET and HEMT Models," Submitted for Publication to the IEEE Transactions on Microwave Theory and Techniques.

International Conference Publications:

1. C. van Niekerk, D. Schreurs, P. Meyer, "Recent Developments in Nonlinear Device Modelling Techniques," IEEE AFRICON '99 Conference, Cape Town, South-Africa, October 1999, pp.1105-1110.

Content

Chapter 1
Introduction

1.1 Introduction

The following sections and chapters will present the results of a study for determining small-signal models for GaAs *metal-semiconductor field-effect-transistors* (MESFET) and *high-electron-mobility-transistor* (HEMT) devices. These transistors are commonly used in MMIC circuits. MMICs offer manufacturers the opportunity to produce microwave systems more cost effectively. However, due to the expense of setting up an MMIC foundry, smaller users have to make use of foundry services provided by external companies [1]. Unlike hybrid MIC circuits, MMICs cannot be tuned after manufacturing, making it necessary to have accurate small-signal and nonlinear device models.

There are several approaches to nonlinear modelling, ranging from equivalent circuit and black box models, to physics-based models that require the solution of the electron transport equations. The physics-based models have the advantage of being scalable since they employ a physical description of the transistor. However, this description of the device requires information that foundries might be unwilling to make available to outside companies. The computation of the model response is also time consuming. This may be problematic, since the iterative methods used in commercial CAD packages calls for hundreds of solutions at a time. While a lot of work is being done on fast physics-based models and their inclusion into CAD software [2,3,4], it is not expected that users outside of device manufacturers will be able to create these models in the near future.

Black box models [5,6,7] have the advantage that they do not require the establishment of links between the device physics and the measured data. They are fast to execute, but are normally limited in the level of nonlinear behaviour that they can represent [6]. Black box models are not commonly found in commercial CAD packages.

Equivalent circuit models are the most common type of models in CAD software. They are constructed from DC measurements and small-signal models extracted from s-parameters measured at different bias points. This approach to nonlinear modelling was first suggested by Willing [8]. Originally, nonlinear models used the extracted data to construct nonlinear functions for the charge and current behaviour of the model according to a first order description

of the device physics [9-16]. Newer models use look-up tables [17,18,19] or neural networks [20] and transforms of the extracted data [21,22,23,24] that remove some of the links that the model has with the device physics. These models can now be called *grey* [18], and have the advantage that their computational evaluation can be performed efficiently in CAD software. Due to the historical impetus behind them, most designers are comfortable with their use and they are easily integrated into existing approximate design techniques [25].

The measurement equipment needed for creating equivalent circuit models is widely available and very little information regarding the device physics is needed. The problem of the model generation is shifted to the software that has to extract and manipulate the model parameters. These factors make this approach very attractive to users of foundry services, or when discrete packaged devices are used.

A key step in the construction of these models is the accurate extraction of the small-signal circuit models from measured s-parameters. This thesis will present new algorithms for performing this task.

1.2 The MESFET and HEMT Small-signal Model

The following paragraphs provide a brief introduction to the GaAs MESFET and HEMT small-signal equivalent circuit model.

The model is divided into an intrinsic and an extrinsic section. The intrinsic part describes the physics of the device, and the values of these circuit elements are a function of the applied bias. The extrinsic elements represent parasitic elements that surround the transistor. They are the result of the transistor packaging and the layout of the device and are considered bias independent.

Figure 1.1 contains the layout of the MESFET and shows the physical origin of the small-signal model elements. The HEMT has a different channel structure, but the model and the function of the small-signal model elements are the same as for the MESFET [26,27].

The grey area under the gate of the transistor represents the depletion region formed in the channel by the gate Schottky diode. The intrinsic capacitors Cgs and Cgd model the distribution of the charge depletion layer beneath the gate. It takes a finite time for the charge depletion layer to change with respect to changes in the gate voltage. This effect is represented by Ri and Rj and

the time constants they form with Cgs and Cgd. The effect of Rj is normally only noticeable at mm-wave frequencies and it is therefore not included in most models. The gain of the device is represented by the transconductance gm, which models the incremental change in I_{ds} for a given change in V_{gs}. Mathematically gm is defined as:

$$g_m = \left.\frac{\partial I_{ds}}{\partial V_{gs}}\right|_{V_{ds}=constant} \tag{1.1}$$

The transconductance cannot respond instantaneously to changes in V_{gs}, and this delay is modelled by τ. The output conductance g_{ds} represents the change in I_{ds} due to changes in V_{ds} and it is described by the following relationship:

$$g_{ds} = \frac{1}{R_{ds}} = \left.\frac{\partial I_{ds}}{\partial V_{ds}}\right|_{V_{gs}\ =\ constant} \tag{1.2}$$

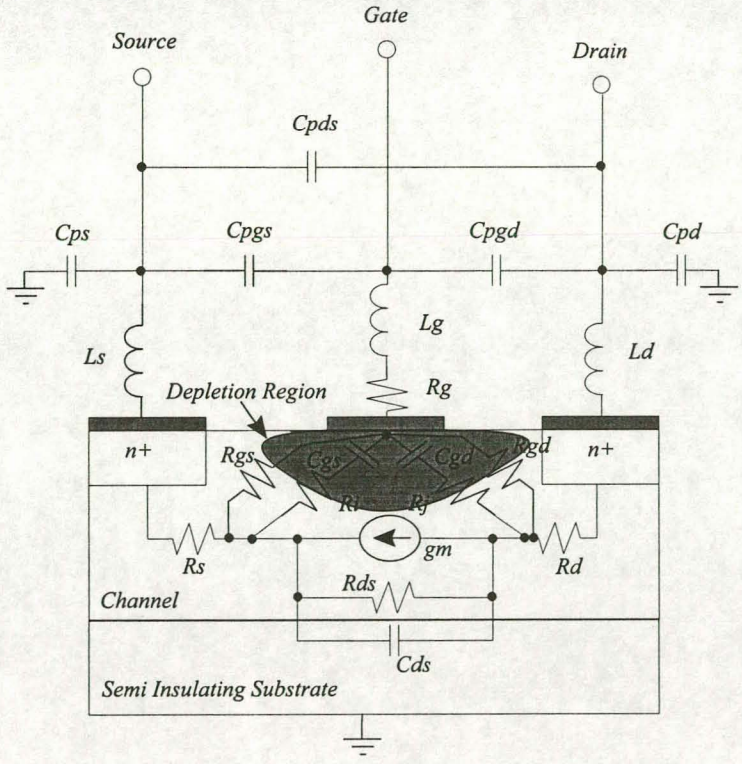


Fig. 1.1 The physical layout of the GaAs MESFET and the origin of the small-signal model elements.

The resistances Rgs and Rgd model the current through the gate Schottky contact [18,27]. Since these currents are normally small, Rgs and Rgd are ignored.

The parasitic gate resistance Rg is primarily determined by the resistance of the Schottky gate contact, while Rd and Rs model the drain and source contact resistances as well as the resistance of the bulk GaAs between the contact and the channel [27,28]. The depletion layer changes with bias, affecting the length of bulk GaAs between the channel and contact. The depletion layer can extend well into the region between the gate and drain for high V_{DS} voltages, thus primarily influencing Rd. At high V_{DS} voltages it has also been predicted that Rd might assume negative values because of the formation of a localised Gunn domain [27,28]. The bias dependencies of Rs and Rd should however be small and they are taken to be bias independent.

The rest of the extrinsic elements, Lg, Ld, Ls and the parasitic capacitors, are determined by the metallisation and layout of the access lines to the device. Not all the parasitic capacitors shown in figure 1.1 are included in the final small-signal model. It is left to the few extrinsic capacitors and the intrinsic capacitors in the final model to absorb their effect. Figure 1.2 shows a popular version of the small-signal model, while figure 1.3 contains two variations that are used when parasitic capacitances cannot be ignored. These are the models most frequently found in the literature and they are also the models used in the rest of this study.

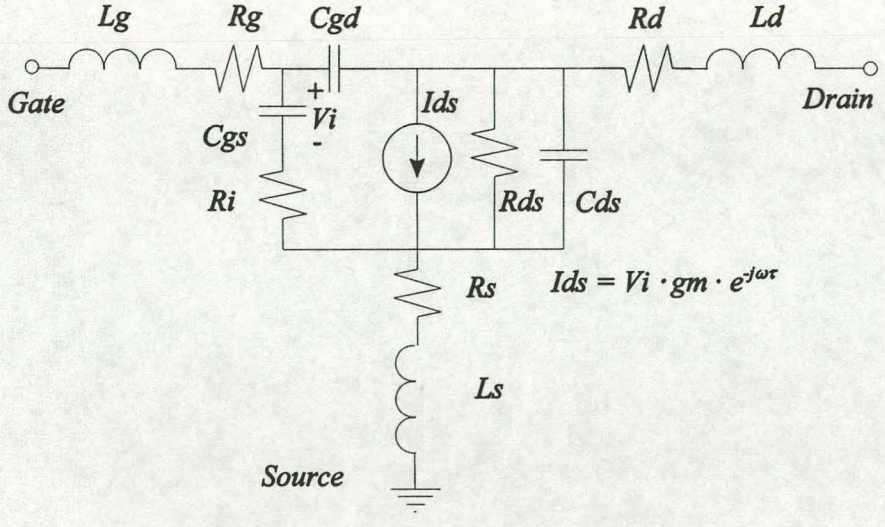


Fig. 1.2 The 13 element FET model.

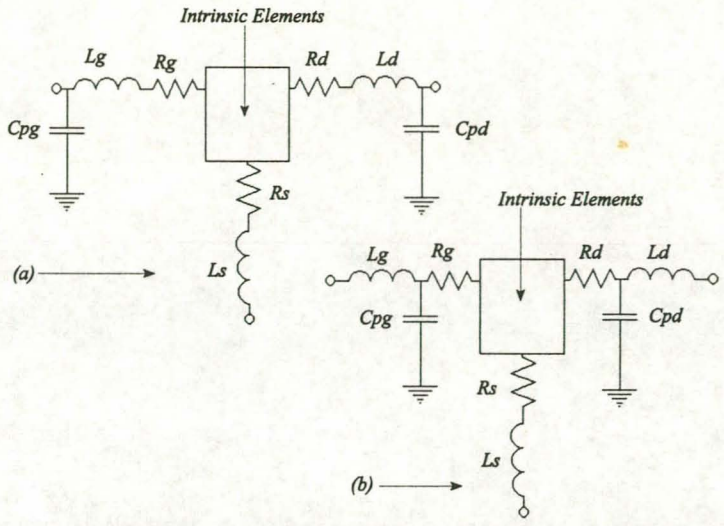


Fig. 1.3 Two variations of the FET model with extrinsic parasitic pad capacitances.

1.3 The Problems Inherent to Extracting Small-Signal Models

Small-signal models are constructed to represent physical processes that occur inside the device. This does not necessarily lead to models that are easy to extract, as correlations exist between the model elements. For instance, due to their position, the gate resistance Rg and the channel resistance Ri look electrically the same, making it difficult to extract their individual values [29]. The model elements can be divided into dominant and non-dominant elements. The dominant elements have a large effect on the predicted s-parameters and their values can therefore be determined accurately regardless of the specific method used. The values of the non-dominant elements can however be varied by large amounts without causing much change in the predicted s-parameters. This is commonly referred to as the uniqueness problem of the extraction [29]. The result is that there are a large number of solutions that appear to provide equally good representations of the measured s-parameters. In order to choose the best of all the possible solutions, it becomes necessary to look at the requirements of the nonlinear model that is to be constructed from the element values. To successfully use the extracted model elements for constructing nonlinear models, the variation of the intrinsic model elements with bias must be smooth. A physically correct solution is therefore required, not just one that provides a good fit to the measured data.

The uniqueness problem is further compounded by the large number of variables that have to be extracted from a limited amount of data. This is commonly referred to as the *curse of dimensions*. The more complex the model is made, the better the fit between modelled and

measured data. It will however become difficult, or even impossible, to determine a unique value for each of the model elements. A balance therefore needs to be found between the model complexity and the accuracy with which the measured data can be represented [30].

There are two different approaches to extracting small-signal models, namely direct extraction and optimiser-based extraction. These will be discussed in more detail in the following sections. The basic problems of small-signal model extraction affect both, although not always in the same way.

1.4 Direct Parameter Extraction

Dambrine [31] was the first author to propose a direct extraction algorithm. He built his method on observations made by Diamand and Laviron [32], i.e. that a cold biased FET (V_{DS} = 0 Volt) has a simpler intrinsic circuit, and therefore presents an alternative way of calculating the extrinsic model elements. Dambrine uses two different cold biases and corresponding simplifications in the model to calculate the values of all the extrinsic elements.

The device is first biased at V_{DS} = 0 Volt and V_{GS} at a suitably large forward bias. In this configuration, the channel of the device is completely open (very few capacitive effects) and the extrinsic inductors and resistors can be determined. At the second cold bias, V_{GS} is held at a voltage less than the pinch-off voltage. The channel conductivity is suppressed, allowing the parasitic capacitors Cpg and Cpd (see figure 1.3) to be calculated. The equations used for determining Cpg and Cpd are valid at the lower measurement frequencies where the parasitic inductors and resistors do not have a large effect.

The direct parameter extraction method has been the subject of considerable discussion and further work [32-57]. Dambrine's original equations for calculating the intrinsic elements were only valid for frequencies below 5 GHz, and more accurate formulas, which are valid for use with broad-band measurements, have been published by Berroth and Bosch [36]. These equations were derived with the aid of the model shown in figure 1.2. Vickes [48] published equations for calculating the intrinsic elements of the model suggested by Curtice [9]. Rorsman [45] proposed a set of equations in which the intrinsic model of figure 1.2 has been extended with the inclusion of the feedback resistance Rj in series with Cgd (see figure 1.1). This model is frequently used to model HEMT devices at mm-wave frequencies.

Much work has been directed at refining and improving the models and methods used to

determine the extrinsic elements [33-35,37,38,42-44,46,47,49,50,52-57]. Dambrine's proposed model for the cold MESFET with the gate held below the pinch-off voltage has been questioned. Several authors found that it over-estimates the value of Cpd when applied to symmetric devices [41,49] and improved cold FET models have been proposed [33,46,49]. Anholt and Swirhun [33] did extensive experimental and theoretical studies concerning the formation of the parasitic capacitors and show that their values are not just dependent on the extrinsic networks, but are also influenced by capacitive effects in the active region of the device. The result of this is that the direct extraction method cannot separate the values of all intrinsic and extrinsic capacitors, specifically Cds and Cpd [33,35]. It is therefore normally assumed that the device is symmetrical in layout and that Cpg = Cpd [39,45], or a ratio of Cpg to Cpd is arbitrarily chosen [47].

The cold bias with the gate forward biased has also been under considerable investigation [34,37,38,41,47,50-57], with different models being proposed. A large number of these models contain current dependent constants whose values are not always certain [47]. The forward bias used for MESFETs cannot be applied when studying HEMT devices due to the formation of a parasitic MESFET inside the HEMT. Various schemes for adapting the direct extraction for HEMTs have been proposed [18,40,41,45,47]. Large forward gate biases and the resultant gate currents also introduce the danger of causing metal migration and other reliability problems [47]. Eskandarian [38] further shows that new expressions are needed for the determining the extrinsic inductors when the parasitic capacitance values become so large that they can no longer be ignored when the device is forward biased.

The direct extraction method calculates a set of model elements for each measurement frequency. The elements are supposed to be frequency independent, but variations do occur because of measurement errors and the approximations used to calculate the element values. The average element value over a certain frequency range is therefore calculated. A certain amount of experience is needed in choosing the best frequencies for determining the different model elements, and the averaging process adds further uncertainty concerning the optimum values of the model elements.

The ill-conditioned nature of the model extraction problem can be seen when the values of the non-dominant elements are evaluated. They exhibit large changes in value as a function of frequency and even nonphysical negative element values are obtained [38,47]. The extrinsic resistances and the intrinsic elements Ri, τ and Cds are especially sensitive. The calculation of the intrinsic elements is further complicated by any errors made in the determination of the dominant extrinsic elements, which leads to a propagation of errors.

The advantage of the direct extraction algorithm is its speed and simplicity. It is also convenient since it can be implemented using only network analyser measurements.

1.5 Optimiser-Based Parameter Extraction

Optimiser-based parameter extraction employs a numerical search technique to minimise an objective function. This objective function represents the difference (error) between the predicted measurements and the measured data obtained from the device. The optimisation algorithm minimises the objective function by iteratively searching for the set of model elements that will result in a function minimum. This was the technique used by Willing [8] when he first proposed the construction of nonlinear models using bias dependent small-signal models.

Optimisation algorithms are also affected by the phenomena described in section 1.3. Due to the correlations between the variables and occurrence of dominant and non-dominant elements in the same problem, the shape of the objective function becomes flat and featureless in the area of the final solution. This makes it difficult to accurately find the function minimum. The large number of variables that have to be extracted also increases the number of local minima in the function.

Both gradient-based optimisation procedures [29,58-70] and random searches [71,72,73] have been used in extraction algorithms. Gradient searches look for points where the first derivative is zero, making them susceptible to terminating in a local minimum. The initial values for a gradient search therefore have to be chosen close enough to the global minimum. A further problem that is encountered by gradient searches, is that at some point, the algorithm will require matrix inversion or the solving of simultaneous equations in order to calculate the next set of improved model element values. Due to the occurrence of dominant and non-dominant elements in the same system, this procedure is numerically ill-conditioned. This causes the non-dominant element values to vary wildly due to numerical noise and other small imprecisions. Random optimisation algorithms are not hampered in this way by local minima, and guided random walks, such as simulated annealing [71,74] and genetic algorithms [72,73,75], can be considered to be quasi globally convergent. They are however very inefficient, especially for a large number of variables, and are only guaranteed to find the global minimum within an infinite number of moves [76]. Standard gradient searches, such as found in commercial circuit simulators, or even random searches, do not provide attractive solutions to the extraction problem.

In recent years there have been several suggestions for new algorithms able to overcome most of the problems associated with optimiser-based extractions [29,58,64,66,67,68,71,77]. These algorithms can be classed as those that strive to reduce the number of variables in the optimisation process [66,67], and methods that try to more effectively deal with large optimisation problems [29,58,64,68,71,77]. The idea of reducing the number of variables is in part based on the direct extraction algorithm approach in that only the extrinsic elements are optimised, and the intrinsic elements are calculated as a function of the extrinsic elements. Lin and Kompa [66] and Shirikawa [67] proposed two methods using this approach. Shirikawa [67] uses the standard equations published by Berroth and Bosch [36] to calculate the intrinsic model elements. During the formulation of his objective function, Shirikawa not only uses the difference between the modelled and measured s-parameters, but also looks at the variance of the calculated intrinsic element values with frequency. A correct choice of extrinsic elements will also minimise the variation in the intrinsic elements as a function of frequency. Lin [66] improves on the method in which intrinsic elements are calculated, by illustrating a new way of deriving a set of robust closed form equations for determining the intrinsic elements. He also provides detailed results indicating the robustness of the method.

Methods have also been developed to deal with large optimisation problems. Kondoh [64] suggested a decomposition approach, and it has been expanded on by Bandler [77]. Kondoh's algorithm has also been used as the basis of hybrid methods that employ his method together with a simulated annealing random search [65]. Patterson [29] showed that the ill-conditioned nature of the optimisation problem can be improved using a systematic analysis employing a principal components sensitivity analysis.

The newer optimisation algorithms are more adapted to the problems of small-signal model extraction and have largely overcome the problem of local minima. These algorithms are considered more accurate than direct extraction methods because they produce a set of model elements that optimally fit the measured data. Optimisation has the further advantage of being a more general approach, making it easier to adapt to new models and technologies. Its primary disadvantages are that it is initially more complex to implement and that it requires more computer resources to solve.

1.6 The Scope of this Study

This dissertation investigates new optimiser algorithms for extracting FET small-signal models from measured s-parameter data. Several contributions are made to the fields of optimiser-based experimental device modelling and optimisation. The dissertation presents the following major contribution [78]:

- A new integrated multi-bias parameter extraction routine is proposed and subjected to rigorous evaluations using data from different GaAs devices.

The new algorithm has the following advantages over standard single bias techniques:

- The new multi-bias algorithm is formulated in such a way that data from different bias points are included into one extraction problem that determines all the elements simultaneously. This allows for more accurate results than multi-bias methods based on single bias or direct extractions, since there is more information available for determining the values of the extrinsic elements. The advantage of this is that the algorithm is less susceptible to bad measurement points.
- Unlike other integrated multi-bias algorithms, which make use of conventional gradient searches [58,69,70], or hybrid gradient and guided random walk procedures [73], the new algorithm efficiently solves the extraction problem using a decomposition-based optimiser.
- The algorithm can achieve lower modelling errors than procedures that exclusively make use of analytical equations to calculate the intrinsic model elements. The lower modelling errors are the result of the freedom that the algorithm has to independently vary the values of the intrinsic model elements.
- The decomposition-based search is shown to be accurate and independent of the choice of optimisation starting values [78]. The accuracy and robustness of the decomposition-based search results from the fact that the algorithm does not make use of one global error function. This makes it more immune to local minima, and the ill-conditioned nature of the problem, than other methods.
- The new procedure makes no assumptions concerning the bias dependencies of non-dominant intrinsic elements [73], nor does the algorithm contain anything that will suppress the bias dependent behaviour of intrinsic elements with a weak dependence on bias [58].
- For the first time it is shown how the accuracy of the extraction results increases as the number of bias points in the integrated extraction process is increased [78].

- The proposed multi-bias algorithm has the further advantage that its formulation is flexible in the sense that specific extrinsic elements can be defined to be bias dependent, should this become necessary.
- Since no assumptions are made in the algorithm concerning the device layout, the algorithm is suitable for modelling commercial devices about which very little additional information is known.

This dissertation also makes the following secondary contributions, mainly concerning the decomposition-based optimisation algorithms that were used [68,79]:

- A new adaptive decomposition-based optimiser that makes use of the principal components sensitivity analysis is presented.
- The factors that influence the principal components sensitivity analysis are discussed in order to provide insight into the way in which the optimisation procedure functions.
- The convergence behaviour of the decomposition-based optimiser is investigated.

The decomposition-based optimiser functions by breaking the problem into sub-problems and solving these problems in a specific sequence. The sequence in which the sub-problems are solved is of key importance to the accuracy and robustness of the decomposition-based optimiser. Previous publications using decomposition either determined the optimisation sequence using experimentation [64], or makes use of algorithms containing a large number of heuristic choices [77].

The new decomposition algorithm, described in chapter 2, uses a principal components sensitivity analysis to determine the optimisation sequence. The procedure analyses the shape of the objective function in order to determine the influence of the different model elements on the error function. This information is used to construct an optimisation sequence that can better handle the problems associated with large ill-conditioned optimisation problems. The proposed algorithm is simple and contains very few heuristic choices, none of which has a large influence on the performance of the algorithm [68,79].

The convergence behaviour of a decomposition-based optimiser is not simple and has not been previously discussed in publications describing the use of such optimisers for large microwave problems [79].

Results are presented on the convergence behaviour of the decomposition-based optimiser, and show that the convergence of the algorithm can be divided into several different phases.

Understanding the convergence behaviour of the algorithm is of crucial importance when the termination conditions of the algorithm are to be specified.

The above mentioned discussions are supported by extensive tests using both simulated and measured data for different GaAs FET devices. These tests illustrate the robustness and accuracy of the algorithms. The test results were also used to confirm the behaviour of the procedure under various conditions.

1.7 Layout of Dissertation

The dissertation has the following layout. Chapter 2 provides a detailed discussion on the decomposition-based optimiser used in this work, as well as its application to single bias extractions. The chapter starts by providing a definition of the decomposition process and the form of the objective function that is to be minimised. The different steps used in the decomposition algorithm are presented in order to develop a deeper understanding of the procedure. These discussions are supported with a variety of experimental results, using both simulated and measured data. The chapter is concluded with a discussion on the convergence behaviour of decomposition-based optimisers.

In chapter 3, the decomposition-based optimiser is adapted to a new integrated multi-bias parameter extraction problem, that overcomes the problems inherent in the single bias extractions performed in chapter 2. A short theoretical development is provided, followed by extensive experimental results using multi-bias measurements from different GaAs FET devices. It is shown how the accuracy of the procedure increases with the use of more bias points in the extraction process. The obtained results are also compared to other optimiser-based and direct extraction methods.

The dissertation is concluded in chapter 4 with an overview of the results and a discussion of future research and extensions to the proposed extraction algorithms.

Chapter 2
Single Bias Decomposition-Based Parameter Extraction

2.1. Introduction

This chapter will describe the development and testing of an adaptive decomposition-based optimiser for FET small-signal model extraction problems. In decomposition, a problem is broken up into smaller problems, which are solved in a specific sequence in order to obtain a solution for the original problem. Decomposition was first used by Kondoh [64] for the extraction of an MESFET model from measured data. Bandler and Zhang [77] did some work on an automatic decomposition algorithm and tried to establish a theoretical basis for describing the technique. They verified Kondoh's work, which is largely based on experimentation, and successfully applied the approach to other optimisation problems.

In most of the mathematical texts concerning the use of decomposition [80] the problems studied can be broken up naturally into smaller parts. The global problem thus consists of independent subproblems that are identified by the decomposition algorithm. In decomposition as used in microwave engineering [64,77], the problem does not consist of linearly independent subproblems. The global problem is broken into subfunctions and it is determined how strong or weak the dependence of these functions are on the different model parameters. The problem therefore has to be broken up along its weakest links. Due to the influence that the subproblems will have on each other, they have to be solved in a specific sequence to ensure the stability of the optimisation algorithm.

For the purpose of this work, the following definition is used when referring to decomposition:

Definition 2.1:

> *Decomposition is a process by which a function that is to be optimised is broken up into several subfunctions. The independent variables of the function are partitioned into groups according to their influence on a particular subfunction. Should the i-th variable have its largest influence on the j-th subfunction, it is assigned to that function. The subfunctions are optimised in a specific sequence, and only with respect to the variables assigned to them. This order is repeated until the whole problem has converged to its final value.*

Section 2.2 describes the objective function that is to be minimised with the decomposition-based approach, while sections 2.3 and 2.4 present a theoretical description of the decomposition-based optimisation algorithm. The chapter is concluded with experimental results using both simulated measurements and measured data for two different GaAs FET devices.

2.2 The Objective Function

In order to extract the model elements, a global objective function must be defined. This error function represents the difference between the measured data and the model response and is minimised by choosing a suitable set of model elements. In the FET extraction there are four different s-parameter measurements that have to be incorporated into the objective function. In defining the global objective function the following criteria must be satisfied:

- No special emphasis may be placed on either the phase or the amplitude of the measured and modelled data.
- No s-parameter may be allowed to dominate the function due to its inherent magnitude.

The objective function is chosen as the sum of the square of errors. Eq. (2.1) presents the function that is to be minimised.

$$F(\bar{x}) = \sum_{i=1}^{N} e_i(\bar{x})^2 \tag{2.1}$$

$$e_i(\bar{x}) = \sum_{j=1}^{2}\sum_{k=1}^{2} \frac{1}{\sigma_{jk}} |R_{jk}(\bar{x},\omega_i) - S_{jk}(\omega_i)| \tag{2.2}$$

$$\sigma_{jk} = |S_{jk}|_{MAX} \tag{2.3}$$

where $F(\bar{x})$ is the global objective function,
$e_i(\bar{x})$ is the magnitude of the error vector at each measured frequency,
$R_{jk}(\bar{x},\omega_i)$ is the model predicted s-parameter at frequency ω_i,
$S_{jk}(\omega_i)$ is the measured s-parameter at frequency ω_i,
$\bar{x}$ is the vector containing the model elements that are to be extracted,
N is the number of frequency points at which s-parameters are measured,
j, k represents the subscripts identifying the four different s-parameters and
σ_{jk} is a normalisation constant equal to the magnitude of the largest s-parameter.

There are other forms of the objective function that will also satisfy the criteria outlined in the beginning of this section, but as will be shown in subsequent sections, the form of eq. (2.1) is convenient since it allows us to easily approximate its second order derivative using only first derivatives, and special methods have been developed for analysing and minimising it [81].

2.3 Subfunctions and Model Element Assignments

In a general optimisation problem, the subfunctions into which the global objective is to be divided, would represent the different system specifications that must be satisfied, for example insertion loss, gain, input and output matching. In the FET model extraction problem, the objective function consists of four subfunctions, namely the errors made by the model in representing the individual s-parameters. Eq. (2.4) represents the general form of the subfunctions into which the global objective function is to be divided.

$$f_{jk}(x_l,\bar{x}_R) = \sum_{i=1}^{N} |R_{jk}(x_l,\bar{x}_R,\omega_i)-S_{jk}(\omega_i)|^2 \qquad (2.4)$$

The definitions of the different components of equation (2.4) are the same as given in section 2.2, with x_l the element that is associated with the subfunction and used to minimise the function. The vector $\bar{x}_R$ represents the other model elements which are kept constant for this function. Note that eq. (2.4) contains no normalisation constants since, unlike the global objective function, there are no s-parameters with largely varying magnitudes competing with each other.

Once the subfunctions have been defined, the next step is to determine with which subfunctions the different model elements are to be associated. This is one of the key steps of the

decomposition algorithm.

In assigning model elements to subfunctions, the sensitivity of the function to an element is the deciding factor, not the magnitude of the subfunctions. Each model element is therefore assigned to the subfunction which has the largest first derivative with respect to that model element. Bandler and Zhang [77] presented the most general formulation for this analysis to date. A sensitivity matrix **Z** is calculated using M randomly chosen sets of model elements in eq. (2.5)

$$Z_{lb} = \sum_{c=1}^{M} \left(\frac{\partial f_b(x_l^c)}{\partial x_l} \right)^2 \frac{x_l^0}{f_b^0} \qquad (2.5)$$

where Z_{lb} is an entry in the n × p matrix **Z**,

$f_b(\bar{x})$ is the b-th subfunction (has the same form as defined in eq. 2.4),

x_l^c is l-th model element at random point c,

x_l^0 and f_b^0 are normalisation constants,

l=1,2,...,n where n is the number of model elements,

b=1,2,...,p where p is the number of subfunctions being evaluated and b = 1 to 4 corresponds to jk = (1,1), jk = (2,1), jk = (1,2) and jk = (2,2).

c=1,2,...,M where M is the number of random points at which eq. (2.5) is evaluated.

If entry (l,b) of matrix **Z** is small, then element l has a very small influence on subfunction b. Matrix **Z** is made sparse by defining a threshold value and setting all entries in **Z** that is smaller than the threshold equal to zero. It now becomes easy to use this matrix to assign model elements to the different subfunctions. Bandler and Zhang present a complete description of the steps necessary to do an automatic decomposition of the model elements [77].

This approach does present problems when dealing with ill-conditioned systems. Because the analysis is performed at random parameter sets, the more dominant model elements will influence the sensitivity values calculated for the less dominant elements, leading to non-optimal element/subfunction associations.

To overcome this, an improved method is presented here. The method uses a number of test problems to which the solution is known. The data describing the test problems are simulated measurements calculated using a small-signal model. Instead of choosing random element sets, all the elements except one (element x_l) are kept at their correct values and x_l is swept through

a range of M points around its correct value. Eq. (2.5) is now used to calculate the entries for the sensitivity matrix. This provides more accurate element/subfunction associations, but it makes the analysis less general.

It is important that there is a good spread of model element/subfunction associations. Assigning too many model elements to the same subfunction will lead to inaccurate results. The decomposition of the elements of the general FET model (see figures 1.5 and 1.6) is shown in table 2.1.

Table 2.1 The decomposition of the general 16 element FET Model

	FET Model Parameter	S-Parameter Subfunction
1	Cgs, Ri, Lg, Rg, Cpg	$f_{11}(x_I,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{11}(x_I,\bar{x}_R,\omega_i)-S_{11}(\omega_i)\rvert^2$
2	Cgd, Rj, Ls, Rs	$f_{12}(x_I,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{12}(x_I,\bar{x}_R,\omega_i)-S_{12}(\omega_i)\rvert^2$
3	Cds, Rds, Ld, Rd, Cpd	$f_{22}(x_I,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{22}(x_I,\bar{x}_R,\omega_i)-S_{22}(\omega_i)\rvert^2$
4	gm, τ	$f_{21}(x_I,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{21}(x_I,\bar{x}_R,\omega_i)-S_{21}(\omega_i)\rvert^2$

In Kondoh's FET parameter extraction algorithm, some of the variables that are assigned to the same subfunction are grouped together, and this group as a whole is optimised. Since some of these groups contain both dominant and non-dominant elements, the accuracy of the procedure may be compromised if the sub-optimisation problems become numerically ill-conditioned.

In the procedure proposed in the following sections, the maximum amount of decomposition possible is applied to the FET parameter extraction problem. Each model element is optimised on its own with respect to the s-parameter on which it has the largest influence. This has the advantage over other techniques of decreasing the correlation that exist between the model elements, and it prevents the different optimisation problems from becoming numerically ill-conditioned.

2.4 Determining the Order of Optimisation

By defining subfunctions and associating the different model elements with specific subfunctions, the optimisation problem is broken up into a number of sub-optimisation problems. The problem is broken along its weakest links in order to decouple the sub-optimisation problems. However, the different subfunctions cannot be decoupled completely, and therefore will not be linearly independent. This makes it important to iteratively step through the sub-optimisation problems in a specific sequence to ensure the convergence of the process. Kondoh [64] determined a fixed order of optimisation through experimentation, while Bandler and Zhang [77] presented a more general algorithm for automatically computing an optimisation sequence. Their algorithm assigns a priority number to each sub-problem according to criteria such as complexity and the magnitude of the error. The procedure relies on a number of heuristic choices and it is not clear how to reproduce their algorithm.

In the next section a new way of computing the order of optimisation using a principal components sensitivity analysis is presented [68]. The proposed method analyses the shape of the objective function and contains few heuristic choices.

2.4.1 Principal Components Sensitivity Analysis

It has been pointed out in chapter 1 that the objective function that is being minimised, is not equally sensitive to all the model elements. In this section it is shown that a principal components analysis can be used to arrange the model elements from those having the largest influence to those having the least influence on the objective function. The sub-optimisation problems for the dominant elements are optimised first, while those for the least dominant elements are optimised last. The reasoning behind this approach is quite simple. By first optimising the dominant model elements, the algorithm tries to get them close to their correct values quickly, thus providing the less dominant elements with a better opportunity to converge to their correct values.

The development of the principal components sensitivity analysis follows that presented by Patterson [29], who used it to improve the conditioning of a conventional multi-dimensional gradient search. The principal components sensitivity analysis was first proposed by K. Pearson [82] as a means of fitting planes by orthogonal least squares. Today it is commonly used to study correlation structures in multivariate statistical problems [83]. In the field of pattern recognition, the procedure is known as the Karhune-Loeve transform (KL). The KL-transform

forms the basis for a class of feature extraction algorithms and is also used to reduce the dimensions of a system in a least squares way. Chen [84] refers to the KL-transform as the discrete analogue of the principal components sensitivity analysis.

In this problem the principal components analysis is used to study the shape of the objective function. Figure 2.1 shows the contours of an objective function $F(\bar{x})$ for a two-dimensional problem. A conventional sensitivity analysis would indicate that the objective function is equally sensitive to both x_1 and x_2, because it does not take correlations between the model elements into account. The objective function is, however, far more sensitive to changes along the y_1 axis than along the y_2 axis. The small gradient of the landscape along the y_2 axis will make it difficult for a numerical optimiser to determine the function minimum along this axis. The errors made in determining the function minimum will influence x_1 and x_2 since both have large components in the y_2 direction. The more elongated the elliptical contours become in the direction of the y_2 axis, the more ill-conditioned the optimisation problem will be.

Geometrically, the axis y_1 and y_2 are the principal axes of the elliptical contours in figure 2.1 [29,85,86]. The principal axes represents the directions in which the landscape of $F(\bar{x})$ respectively show the largest and least change in its values. This property of the principal axes vectors, namely the orientation, is what sets it apart from other orthogonal basis functions that can be used to describe the system [87].

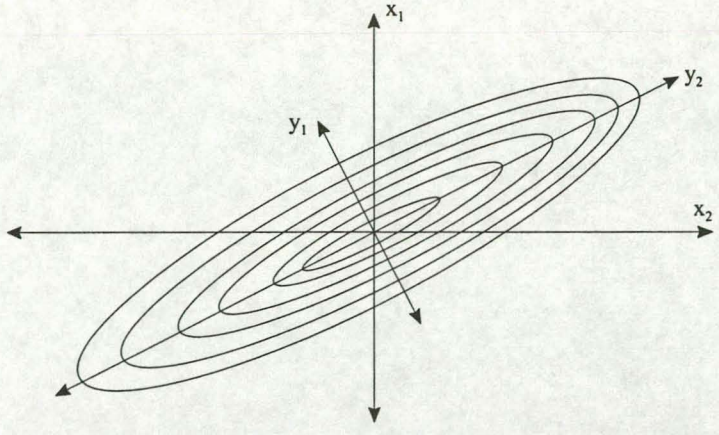


Fig. 2.1 The contours of a two dimensional landscape of an objective function. The contours illustrates the ill-conditioned nature of the problem and the correlations that exist between the variables.

In order to derive the principal axes, the error function $F(\bar{x})$ is expanded in a Taylor series around the optimal point. This series is truncated at the second order term to give eq. (2.6)

$$F(\bar{x}^* + \delta\bar{x}^*) = F(\bar{x}^*) + (\delta\bar{x})^t \nabla F(\bar{x}^*) + \frac{1}{2}(\delta\bar{x})^t \nabla^2 F(\bar{x}^*)\delta\bar{x} \qquad (2.6)$$

where $\bar{x}^*$ is the optimal set of element values at which $F(\bar{x})$ is a minimum, and $\delta\bar{x}$ is a small difference. In the region of the solution the linear terms will be small or zero and the landscape of $F(\bar{x})$ is described primarily by the quadratic term $(\delta\bar{x}^t)\nabla^2 F(\bar{x}^*)\delta\bar{x}$ [29,86]. $\nabla^2 F(\bar{x}^*)$ is also known as the Hessian matrix of $F(\bar{x})$, or **H**.

Equation (2.7) represents the quadratic term in scalar form, with h_{ij} representing the elements of the Hessian matrix.

$$\delta\bar{x}^t \boldsymbol{H} \delta\bar{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} h_{ij}\,\delta x_i\,\delta x_j \qquad (2.7)$$

The Hessian is a real symmetric matrix, which means that its eigenvalues will thus be real and it can be expanded to the form shown in eq. (2.8) [85].

$$\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^t \qquad (2.8)$$

where **Q** is an orthogonal matrix whose columns are the eigenvectors of **H** and
D is a diagonal matrix with the i-th element on the diagonal being equal to the i-th eigenvalue.

A new vector $\bar{y}$ is defined that is related to $\delta\bar{x}$ by the linear transformation given in eq. (2.9).

$$\delta\bar{x} = \boldsymbol{Q}\bar{y} \qquad (2.9)$$

Equation (2.10) shows that the substitution of eq. (2.9) into eq. (2.7) leads to a simpler form of the quadratic in which the cross-product terms in eq. (2.7) have been eliminated [85].

$$\begin{aligned}\delta\bar{x}^t\boldsymbol{H}\delta\bar{x} &= (\boldsymbol{Q}\bar{y})^t\boldsymbol{H}(\boldsymbol{Q}\bar{y}) \\ &= \bar{y}^t\boldsymbol{Q}^t\boldsymbol{H}\boldsymbol{Q}\bar{y} \\ &= \bar{y}^t\boldsymbol{D}\bar{y} \\ &= \sum_{i=1}^{n} d_{ii}y_i^2\end{aligned} \tag{2.10}$$

D is a diagonal matrix containing the eigenvalues of **H** and d_{ii} is the i-th element on the diagonal of **D**.

The simplification achieved in eq. (2.10) is the result of the choice of a new co-ordinate system. Geometrically, the process corresponds with a rotation of the axes shown in figure 2.1 from the x_1, x_2 axes to the y_1, y_2 axes. This rotation is achieved by the linear transformation in eq. (2.9), with the eigenvectors of **H** (columns of **Q**) presenting the principal axes of the system as shown in figure 2.1. The eigenvectors of **H** are known as the principal components of the system [29,85,87].

Because **Q** is an orthogonal matrix, the linear transformation shown in eq. (2.9) will only result in a rotation of the landscape shown in figure 2.1, but not a change of its shape. The eigenvector corresponding to the largest eigenvalue is the direction for which $F(\bar{x})$ will show the maximum sensitivity to changes in $\bar{x}$. The eigenvector corresponding to the smallest eigenvalue is the direction of least sensitivity. This is because the length of the i-th eigenvector is proportional to $\sqrt{d_{ii}}$ [29,87,88]. In section 2.4.3, this analysis will be viewed from a different perspective.

The principal components sensitivity analysis has now been reduced to an analysis of the eigenvalues and eigenvectors of the Hessian matrix $\nabla^2 F(\bar{x}^*)$. To apply this to the problem at hand, we use the objective function defined in section 2.2.

$$F(\bar{x}) = \sum_{i=1}^{N} e_i(\bar{x})^2$$

The partial derivative of this equation with respect to the l-th model element is,

$$\frac{\partial F(\bar{x})}{\partial x_l} = 2\sum_{i=1}^{N} e_i(\bar{x})\frac{\partial e_i(\bar{x})}{\partial x_l} \tag{2.11}$$

Writing eq. (2.11) in vector format to account for all the different model elements leads to,

$$\begin{bmatrix} \frac{\partial F(\bar{x})}{\partial x_1} \\ \vdots \\ \frac{\partial F(\bar{x})}{\partial x_n} \end{bmatrix} = 2 \begin{bmatrix} \frac{\partial e_1(\bar{x})}{\partial x_1} & \cdots & \frac{\partial e_N(\bar{x})}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial e_1(\bar{x})}{\partial x_n} & \cdots & \frac{\partial e_N(\bar{x})}{\partial x_n} \end{bmatrix} \begin{bmatrix} e_1(\bar{x}) \\ \vdots \\ e_n(\bar{x}) \end{bmatrix} = 2\,\boldsymbol{J}^t \bar{e} \tag{2.12}$$

where **J** is the Jacobian matrix of $F(\bar{x})$ and $\bar{e}$ is a vector consisting of the modelling errors made at the N different measurement points. The Jacobian matrix is defined as

$$\boldsymbol{J}(\bar{x}) = \begin{bmatrix} \frac{\partial e_1(\bar{x})}{\partial x_1} & \frac{\partial e_1(\bar{x})}{\partial x_2} & \cdots & \frac{\partial e_1(\bar{x})}{\partial x_n} \\ \frac{\partial e_2(\bar{x})}{\partial x_1} & \frac{\partial e_2(\bar{x})}{\partial x_2} & \cdots & \frac{\partial e_2(\bar{x})}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial e_N(\bar{x})}{\partial x_1} & \frac{\partial e_N(\bar{x})}{\partial x_2} & \cdots & \frac{\partial e_N(\bar{x})}{\partial x_n} \end{bmatrix} \tag{2.13}$$

where

$$e_i(\bar{x}) = \sum_{j=1}^{2} \sum_{k=1}^{2} \frac{1}{\sigma_{jk}} |R_{jk}(\bar{x},\omega_i) - S_{jk}(\omega_i)| \tag{2.14}$$

Differentiating eq. (2.11) again leads to,

$$\frac{\partial^2 F(\bar{x})}{\partial x_l^2} = 2\sum_{i=1}^{N} \frac{\partial e_i(\bar{x})}{\partial x_l} \frac{\partial e_i(\bar{x})}{\partial x_l} + 2\sum_{i=1}^{N} e_i(\bar{x}) \frac{\partial^2 e_i(\bar{x})}{\partial x_l^2} \tag{2.15}$$

$$\frac{\partial^2 F(\bar{x})}{\partial x_l \partial x_r} = 2\sum_{i=1}^{N} \frac{\partial e_i(\bar{x})}{\partial x_l} \frac{\partial e_i(\bar{x})}{\partial x_r} + 2\sum_{i=1}^{N} e_i(\bar{x}) \frac{\partial^2 e_i(\bar{x})}{\partial x_l \partial x_r} \tag{2.16}$$

The Newton approximation [29,89] states that in the region of the solution $e_i(\bar{x})$ will be small, making it possible to neglect the second term in eq. (2.15) and (2.16). Only first order differentials are therefore needed to approximate the Hessian $\nabla^2 F(\bar{x}^*)$. In vector notation, the approximation for the Hessian matrix can be written as,

$$\boldsymbol{H} = 2\boldsymbol{J}^t\boldsymbol{J} \tag{2.17}$$

Table 2.2 The Q matrix of an example calculated using simulated measurements. Only the components contributing more than 1% to the magnitude of the matrix are shown. Beneath each of the eigenvectors, the corresponding eigenvalue is listed.

Parameters	Eigenvectors												
x_1 = Lg				0.078	0.1	0.36	0.93						
x_2 = Rg												0.59	0.81
x_3 = Cgs		0.34	0.94	0.02									
x_4 = Ri												0.81	0.59
x_5 = Cgd	0.99	0.13	0.1										
x_6 = gm											1		
x_7 = τ				0.097	0.99	0.1	0.1						
x_8 = Rds										1			
x_9 = Cds	0.14	0.93	0.34	0.015									
x_{10} = Rd									1				
x_{11} = Ld				0.218	0	0.9	0.37						
x_{12} = Rs								1					
x_{13} = Ls			0	0.968	0.1	0.24							
Eigenvalues	4.97e+29	9.33e+26	5.67e+24	1.90e+21	3.49e+20	2.93e+17	4.91e+15	2.36e+13	1.87e+08	1.08e+03	8.74e+02	5.15e-08	9.50e-09

Table 2.2 shows an example of a **Q** matrix. The corresponding eigenvalue is shown beneath each column. For the sake of clarity, only the vector components that contribute more than 1% to magnitude of each eigenvector is shown. From table 2.2 it can be seen that the direction of the most sensitive principal component is dominated by the intrinsic capacitors Cgs, Cgd and Cds, with the largest contribution coming from Cgd. Cgd is the most dominant element in the system. By starting with the most sensitive principal component and looking at the largest components of the eigenvector, it is possible to order the model elements from those that are the most dominant to those that have the least effect on the objective function. The next section will discuss a simple algorithm for automating this process.

2.4.2 Automating the Sensitivity Analysis

A simple algorithm was developed to calculate the order of optimisation. The Jacobian matrix is calculated using central difference numerical differentiation and the Hessian matrix is determined using the approximation given in eq. (2.17). The eigenvalues and corresponding eigenvectors of the Hessian matrix are calculated using the EISPACK mathematical library [90]. The following steps outline the procedure that is followed to obtain the optimisation sequence from the eigenvectors and eigenvalues of the Hessian matrix.

- Assign starting values to the model elements and calculate **H**.
- Determine all the eigenvalues and eigenvectors of **H**.
- Take the absolute values of the eigenvalues and their corresponding eigenvectors. All that is needed in the sensitivity analysis is magnitude data - unnecessary sign information will only complicate the sorting procedures. The Hessian matrix $2\mathbf{J}^t\mathbf{J}$ is at worst positive semi-definite [89], which means that its eigenvalues should always be positive. The eigenvalues and eigenvectors are calculated with an iterative numerical procedure and are accurate with respect to the norm of the Hessian matrix [90]. This means that for an ill-conditioned system, the smaller eigenvalues may be given as negative. There are two methods for overcoming this problem. The first approach takes the absolute value of all the components, while the second divides the elements of the Hessian with a constant number before performing the eigenvalue analysis. Both methods provide the same results.
- Sort the eigenvalues in ascending order. The eigenvectors are stored as the columns of an eigenvector matrix **Q**. The sorting process changes the position of the columns in matrix **Q** as well.
- Matrix **Q** is made sparse by eliminating all the non-essential entries in each vector. All the components of each column of **Q** that contributes less then λ % to the vector magnitude is set equal to zero.
- Some model elements will have entries in more than one column. The gate source capacitance may for instance be a component in both the eigenvectors corresponding to the largest and the second largest eigenvalue. To ensure that every row of matrix **Q** only has one entry, only the entry in each row corresponding to the largest eigenvalue is retained.
- Should the value of λ be too large, a situation can result in which one of the model elements does not have an entry in the sparse matrix **Q** that can be used to determine its position in the optimisation order. An error check is made to ensure that each row of the sparse matrix **Q** has at least one entry. Should this not be the case, then λ has to be

reduced. More will be said in the next section concerning the choice of λ.

The sparse matrix **Q** is used to sort the model elements in the order of decreasing influence on the objective function. This is done by selecting the largest component in the eigenvector corresponding to the largest eigenvalue. The component's row index will correspond to the most sensitive model element. Once the row index has been found and stored, the component is set equal to zero. This procedure is repeated for all the other components in the vector, and once they have been sorted, for all the other vectors.

Table 2.3 shows an example of the eigenvector matrix **Q** that has been reduced to its sparse form. The table also contains the optimisation sequence that was obtained for the model elements.

Table 2.3 An example of the eigenvector matrix **Q** after it has been reduced to its sparse form. A λ of 10% was used to eliminate elements from the eigenvectors. The number next the model elements in the first column indicates the position of the element in the optimisation sequence.

Parameters and Opt. Sequence		Eigenvectors												
Lg	7						0.359							
Rg	13												0.59	
Cgs	3		0.344											
Ri	12												0.807	
Cgd	1	0.99												
gm	11											1		
τ	6					0.992								
Rds	10										1			
Cds	2	0.14												
Rd	9									1				
Ld	5				0.218									
Rs	8								1					
Ls	4				0.968									
Eigenvalues		5e+29	9e+26	6e+24	2e+21	3e+20	3e+17	5e+15	2e+13	2e+08	1077	874.1	5e-08	1e-08

2.4.3 Factors that Influence the Optimisation Sequence

There are several factors that can influence the calculated optimisation sequence, and thus the accuracy of the extraction procedure. A wide range of experiments were conducted to isolate these factors and determine the sensitivity of the extraction algorithm to them. The following factors were all found to cause small changes in the optimisation sequence.

- **Numerical uncertainty**. There is numerical uncertainty resulting from numerical differentiation, the finite numerical precision of the computers used and the iterative techniques used to calculate the eigenvalues and the eigenvectors. Numerical uncertainty influences the position of the less dominant model elements. There was no evidence that the extraction accuracy is sensitive to these changes.
- **The starting values used in the calculation of H.** The order of optimisation showed a slight dependence on the optimisation starting values. The assumption on which the analysis of section 2.4.1 is built, is that $\bar{x}$ is close to the solution $\bar{x}^*$ of the problem. Since the first optimisation sequence is calculated with the elements far from the solution, it was reasoned that there might be an advantage to recalculating the optimisation sequence periodically. No indications could be found that this improved the extraction accuracy.
- **The influence of the factor λ used for making Q sparse**. Two possible choices can be made for λ. A standard empirical value for λ can be picked that works for all cases and can be considered safe. A λ of 10% would be such a value. All the components of an eigenvector that contribute less than 10% to the magnitude of the vector will be set to zero. The second choice for λ is the largest value for λ that still results in all the model elements having at least one entry in one of the columns of the sparse matrix **Q**. The choice of λ was not found to have a large influence on the order of optimisation. There were however examples in which a high value of λ led to relative large changes in the order of the more dominant model elements, leading to a decrease in the extraction accuracy. A value of 10% was therefore used for λ.

When the optimisation sequences of different examples are compared, a basic pattern emerges. This pattern is further highlighted when the model is extended with the addition of more components. Capacitors are always dominant, followed by τ and the extrinsic inductors. The intrinsic and extrinsic resistors and the transconductance gm are always the least sensitive elements in the model. Table 2.4 presents the optimisation sequence for two different MESFETs as an example.

Table 2.4 Element values for two examples		
Position in Optimisation Sequence	Parameter	
	Transistor 1	Transistor 2
1	Cgd	Cgd
2	Cgs	Cgs
3	Cds	Cds
4	τ	τ
5	Ls	Ls
6	Lg	Lg
7	Rs	Ld
8	Ld	Rs
9	Rd	Rd
10	gm	gm
11	Rds	Rds
12	Ri	Rg
13	Rg	Ri

To gain more insight into this pattern in the optimisation sequence, it becomes necessary to look at how the principal components analysis is used in multivariate statistical analysis and the KL-transform. In the previous sections the technique has been viewed as an analysis of the Hessian matrix, and thus the curvature of $F(\bar{x})$. In multivariate statistical analysis the method is applied to a data matrix **X**, whose columns represent different sets of measurements, each having a zero mean. From the data matrix, the covariance matrix is calculated using

$$\boldsymbol{C}_{XX} = \frac{1}{m-1}\boldsymbol{X}'\boldsymbol{X} \tag{2.18}$$

where $\mathbf{C}_{XX}$ is the covariance matrix,
m is the number of columns in the data matrix and
X is the data matrix of which the columns have zero mean.

The matrix $\mathbf{C}_{XX}$ is then used to investigate the properties of **X** [88]. When eq. (2.17) and (2.18) are compared, it is clear that the covariance matrix and Hessian matrix have the same structure. The principal components analysis can therefore be seen as a study of the Jacobian matrix and

the correlations that exist between the derivatives of the model elements. Relating the sensitivities to the Jacobian matrix has the advantage of allowing a one to one correspondence between the columns of the matrix and the model elements, something which is lost in the calculation of the Hessian matrix.

The eigenvalue equation for the Hessian matrix **H** can be written in matrix form as,

$$2\boldsymbol{J}^t\boldsymbol{J}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{D} \tag{2.19}$$

with **D** being a diagonal matrix containing the eigenvalues of $2\mathbf{J}^t\mathbf{J}$. Since **Q** is an orthogonal matrix, its inverse is equal to its transpose. Pre-multiplying both sides of eq. (2.19) with $\mathbf{Q}^t$ leads to,

$$2\boldsymbol{Q}^t\boldsymbol{J}^t\boldsymbol{J}\boldsymbol{Q} = \boldsymbol{D} \tag{2.20}$$

A new matrix **Y** is defined as,

$$\boldsymbol{Y} = \boldsymbol{J}\boldsymbol{Q} \tag{2.21}$$

which simplifies eq. (2.20) to,

$$2\boldsymbol{Y}^t\boldsymbol{Y} = \boldsymbol{D} \tag{2.22}$$

Y is a new transformed data matrix with the same dimensions as matrix **J**. The columns of **Y** are formed by linear combinations of the rows of the original data matrix **J** and the eigenvectors of the Hessian matrix **H**. Since eq. (2.22) leads to a diagonal matrix **D**, it can be stated that the transformation has also removed any correlation between the columns of **Y**. The eigenvalues on the main diagonal of **D**, can be expressed in terms of the elements of **Y**.

$$\lambda_{ii} = 2\sum_{j=1}^{N} y_{ji}^2 \tag{2.23}$$

where i is the column number and j is the row number of the matrix **Y**. Eq. (2.23) represents the variance of the data in column i of the new data matrix **Y**. The total variance of the data in **Y** is the sum of all the eigenvalues. The total variance is the same as the total variance of the data in the original data matrix **J** [87,88]. This is to be expected since the orthogonal transformations used in the principal components analysis preserves the length and the angle between the

transformed vectors [85]. Since the transformations only perform rotations and translations on the data, but do not change the shape, the variance of the data is preserved.

From the equations listed above, it can be seen that the first principal component is the linear combination of the elements of **J** for which the variance is maximum. The second principal component is uncorrelated with the first, and explains the largest portion of the remaining (residue) variance. The other principal components are defined in a similar way, each being orthogonal to the others and representing the maximum amount of the residual variance [87,88].

Each column of **J** represents the derivative of the modelling error with respect to a different model element at the different measurement frequencies. The orientation of the most important principal components will thus be dominated by the model elements that show the largest variation in their modelling errors as a function of frequency. It is therefore to be expected that the model elements whose transfer functions are themselves a function of frequency, will have the largest influence on the variation of the derivatives in **J**.

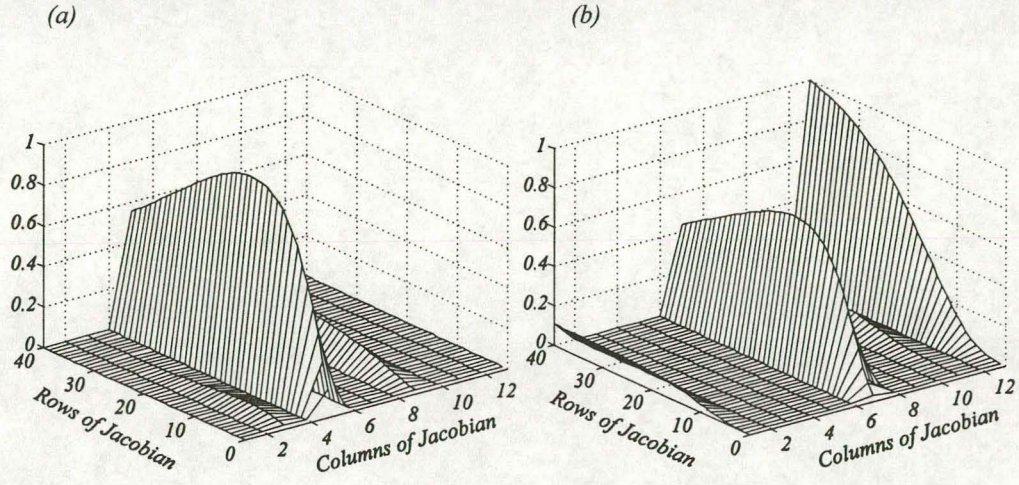


Fig. 2.2 The normalised amplitude of the Jacobian matrix for the FLR016XV device example listed in table 2.5. The columns correspond to the sequence of the model elements in table 2.5. Figure (a) shows the whole matrix, illustrating the dominant columns corresponding to the intrinsic capacitors, with (b) the matrix with these columns set to zero to show the inductors and channel delay τ as the next level of dominant elements.

Figure 2.2 shows the normalised amplitude of a representative Jacobian matrix. It is clear that the elements with the largest frequency dependence correspond to the most prominent columns of the matrix, which will account for their prominence in the sensitivity analysis. In contrast, the importance of the frequency independent elements such as the resistors and the transconductance is more related to their position in the model and their size. The position of a model element in the order of optimisation is therefore determined by the following three factors (in order of importance):

- The frequency dependence of the element value.
- The magnitude of the element value.
- The position of the element in the model.

These three criteria provide rough indicators as to how the sensitivity analysis works and what really influences the order of optimisation. The effect of the starting point of the search is not included in this discussion, but differences in starting values and measured data do have small influences on the order of optimisation. This is because $e_i(\overline{x})$ is a function of both the model element values and the measured data (see eq.(2.2)).

2.5 Experimental Results Using Simulated Measurement

2.5.1 Absolute Accuracy Tests

The FET extraction algorithm was tested using simulated s-parameter data to determine their absolute accuracy and robustness. The data was generated using an FET small-signal model and has the advantage that it contains no unknown imperfections, making it possible to obtain a perfect fit. This allows the behaviour of the algorithms to be studied, without having to contend with factors that are not controllable.

Four different test problems were used. Table 2.5 contains the element values for the 13 element MESFET models used in the tests (see figure 1.5). The element values correspond to those found in the 1993 Fujitsu data book [91]. The s-parameters were generated at 44 frequency points, starting at 1 GHz and ending at 44 GHz.

Table 2.5 Transistor Test Examples				
Parameter	FLR016XV	FLR106XV	FLK052XV	FHR02X
1. Lg (nH)	0.184	0.076	0.14	0.1
2. Rg (Ω)	0.15	0.02	0.36	1.3
3. Cgs (pF)	0.28	2	1.31	0.2
4. Ri (Ω)	5.23	0.693	2	2.5
5. Cgd (pF)	0.0095	0.12	0.065	0.025
6. gm (mS)	24.5	160	80	50
7. τ (psec)	5	5	6	0.85
8. Rds (Ω)	552	60	135	188
9. Cds (pF)	0.0988	0.55	0.36	0.049
10. Rd (Ω)	4	0.7	1	1.3
11. Ld (nH)	0.151	0.069	0.08	0.1
12. Rs (Ω)	1.5	0.2	0.9	1.3
13. Ls (nH)	0.0054	0.0077	0.035	0.08

The robustness of the procedure was tested by performing a large number of extractions from random starting values that were picked throughout a large search space using a uniform distribution. The allowed search space was defined as 0.1 to 5 times the correct element values listed in table 2.5. One hundred extractions were performed for each of the devices.

Table 2.6 provides a summary of the results. The performance of the algorithm is described using the average extracted element values, as well as the minimum and maximum element value extracted for each model element. These terms are represented as an error, expressing the percentage difference between the obtained value and the correct model element value. The table therefore provides both the average error, as well as the full range around the correct element value in which extraction solutions were obtained. This representation is also used in all subsequent tables that summarises the results of robustness tests.

Table 2.6 The error in the extracted model elements. The first column for every transistor represents the average error, while the second and third columns represent the minimum and maximum element values extracted. All the errors are given as a percentage of the correct element value.

	FLR106XV			FLR016XV			FLK052XV			FHR02X		
	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Cgs	0.11	-0.08	0.34	0.02	-0.04	0.16	0.79	-0.56	2.43	0.51	-3.82	4.37
Cgd	0.17	-0.53	0.12	0.04	-0.33	0.08	1.35	-4.2	0.98	0.6	-3.35	3.23
Cds	0.02	-0.02	0.07	0	0	0.03	0.15	-0.1	0.46	0.74	-1.56	5.61
gm	0.1	-0.07	0.29	0.02	-0.03	0.14	0.69	-0.49	2.14	0.06	-3.46	5.3
τ	0.06	-0.07	0.21	0.01	-0.04	0.12	0.55	-0.42	1.71	1.5	-7.48	4.11
Ri	4.76	-14.64	3.3	1.64	-12.92	2.92	28.56	-87.93	20.06	8	-89.91	81.34
Rds	0.09	-0.29	0.07	0.01	-0.14	0.06	0.68	-2.11	0.51	0.41	-4.06	3.47
Lg	0.01	-0.03	0	0	-0.06	0.02	0.09	-0.28	0.09	0	-0.18	0.19
Ld	0.5	-0.01	0.01	0	-0.02	0.01	0.03	-0.04	0.11	1	-4.35	0.45
Ls	0.02	-0.07	0.02	0.03	-0.19	0.06	0.06	-0.19	0.07	0.06	-0.25	0.22
Rg	129.9	-90	400	50.62	-90	398.7	129.8	-90	400	10.25	-89.99	104.8
Rd	0.72	-2.32	0.58	0.16	-1.4	0.35	7.71	-23.79	5.69	23.25	-41.01	95.51
Rs	2.95	-2.053	9.06	0.5	-0.89	3.93	9.47	-6.95	29.12	4.73	-52.06	61.62

Figures 2.3 to 2.6 provide a graphical representation of the robustness test for the FLR106XV. The x-axis of the graphs show the i-th extraction performed, while the y-axis show the value of a specific model element. The solid line depicts the extracted value while circles show the starting value for each element.

Table 2.6 shows that all the model elements, with the exception of the parasitic gate resistance Rg, can be extracted with a high degree of accuracy. Rg is one of the least dominant elements in the model and its value is masked by small errors made in the extraction of the other elements. It can be noted from table 2.6 that not all the extraction examples provide the same level of accuracy. The FHR02X device has a far larger spread of extracted element values than any of the other examples, but it also has a far smaller spread for the extracted Rg values. A simple first order sensitivity analysis indicates that the FHR02X has a higher sensitivity towards Rg than any

of the other devices tested. The different examples that were studied led to the conclusion that the achievable extraction accuracy is linked to the model's dependence on Rg, with a higher dependence leading to more difficult extractions. No conclusive experiments could be constructed that would indicate what causes this dependence.

Kompa and Novotny [63] found that if the value of the parasitic resistance Rg is known, the values of the other model elements can be extracted with high accuracy. Similar results were found for the work presented here. Table 2.7 contains the extraction results for the FLR106XV and FHR02X devices when Rg is fixed at its correct value. A decrease in the extraction errors for all the model elements is observed. It is important to note that by fixing the value of Rg, all the extraction problems do not become equally simple. There is still a difference between the different examples in terms of the accuracy that can be achieved.

Table 2.7 The errors made in extracting model elements when Rg is fixed at its correct value. All the errors are given as a percentage of the correct element values.

	FLR106XV			FHR02X		
Parameter	Avg	Min	Max	Avg	Min	Max
Cgs	0.0022	-0.0088	0.0052	0.0222	-0.108	0.375
Cgd	0.0028	-0.0176	0.0131	0.207	-0.123	0.415
Cds	0.003	-0.0121	0.0108	0.91	-0.339	4.894
gm	0.0011	-0.0017	0.0062	0.372	-0.155	2.219
τ	0.0102	-0.0202	0.0404	1.015	-5.269	0.356
Ri	0.0031	-0.0079	0.0146	2.158	-8.779	0.81
Rds	0.0034	-0.0246	0.0086	0.111	-1.183	0.0891
Lg	0.0009	-0.0072	0.0006	0.0341	-0.141	0.0153
Ld	0.0015	-0.0054	0.0062	0.976	-4.602	0.312
Ls	0.0002	-0.0016	0.0036	0.0047	-0.356	0.0795
Rg	0	0	0	0	0	0
Rd	0.027	-0.1973	0.0762	18.995	-6.481	93.145
Rs	0.007	-0.0391	0.0243	2.75	-1.285	17.274

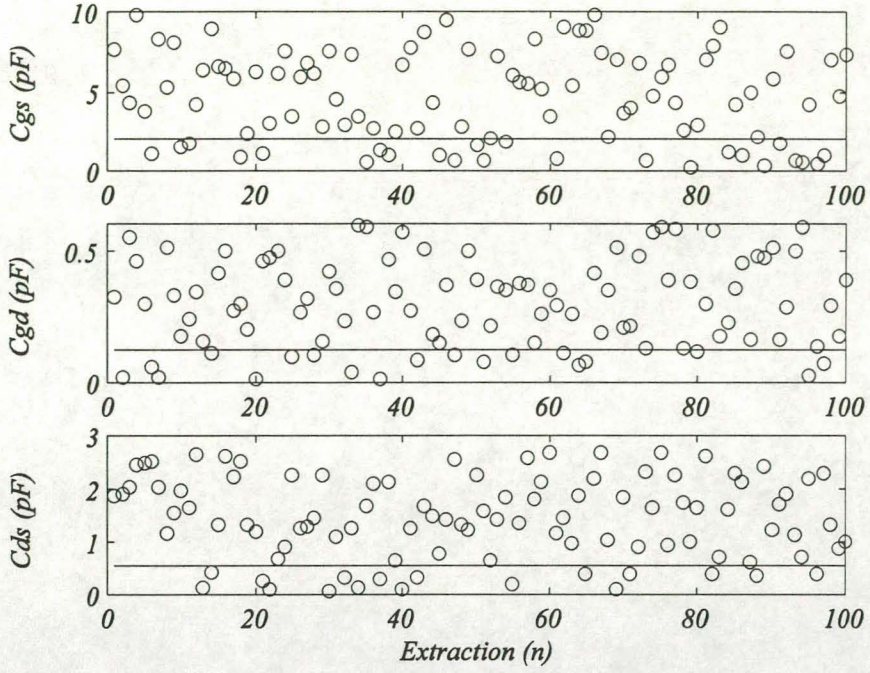


Fig. 2.3 The variation in the intrinsic capacitors for a hundred extractions using random starting values for the FLR106XV transistor. (○ starting value, – final value)

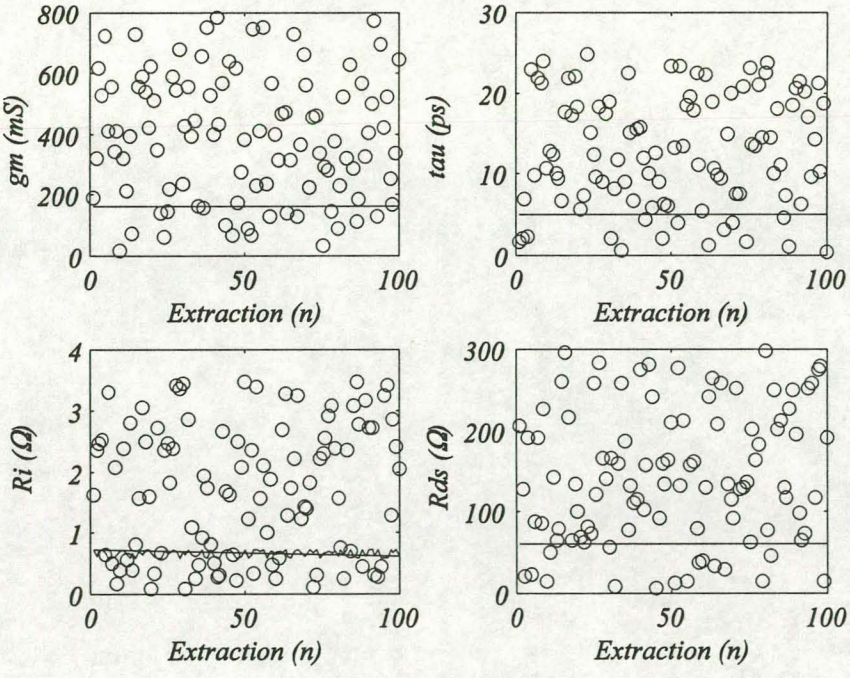


Fig. 2.4 The variation in the intrinsic current source parameters and the intrinsic channel resistances Ri and Rds as seen over a 100 extractions done with random starting values for the FLR106XV transistor. (○ starting value, – final value)

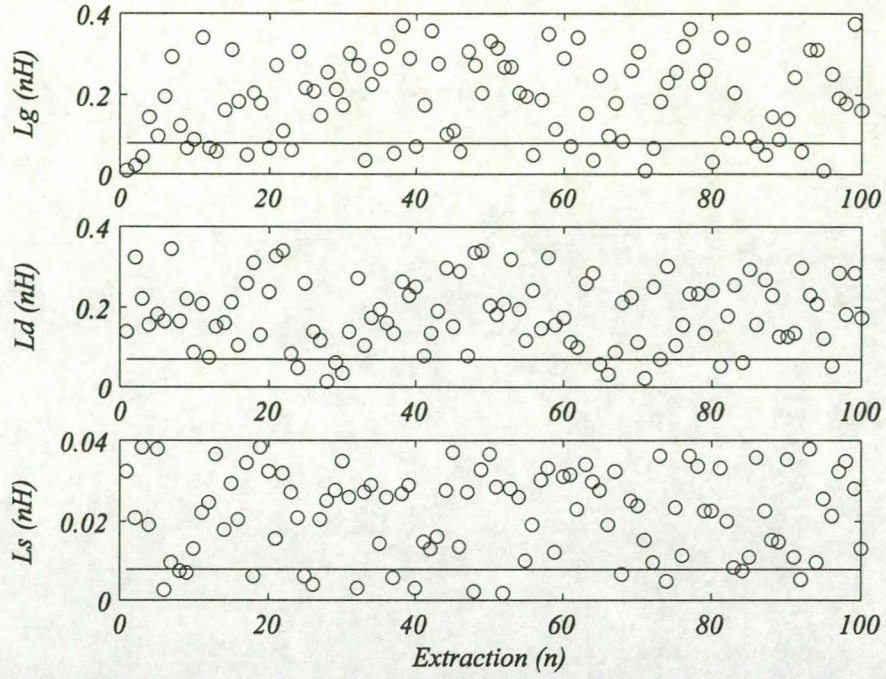


Fig. 2.5 The variation of the extrinsic capacitors when viewed over a 100 extractions with random starting values for the FLR106XV transistor. (○ starting value, – final value)

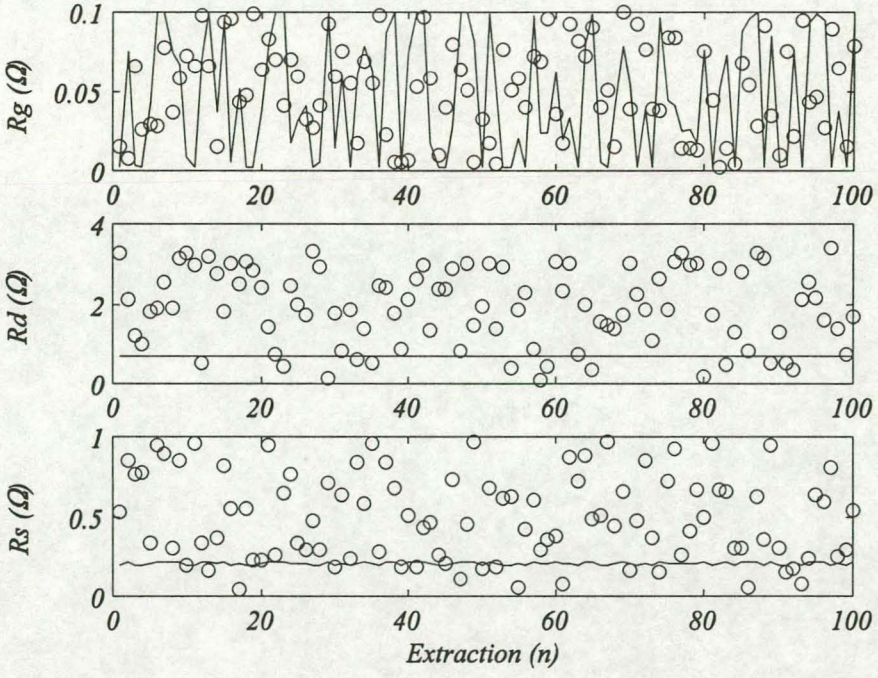


Fig. 2.6 The variation of the parasitic resistances over a 100 extractions started with random starting values for the FLR106XV transistor. (○ starting value, – final value)

The distributions of the extracted model elements were also investigated. An example of a typical distribution is shown in figures 2.7 to 2.9 for the FLK052XV transistor. It should be noted that this distribution does not conform to a standard statistical description. The variations that are seen in the extracted model element values are a result of the strong correlations that exist between the model elements. The interactions between the model elements are dependent on the shape of the objective function, which is not something that can be described statistically. Often used terms such as standard deviation may therefore be misleading if used to describe the accuracy of the extraction process [79].

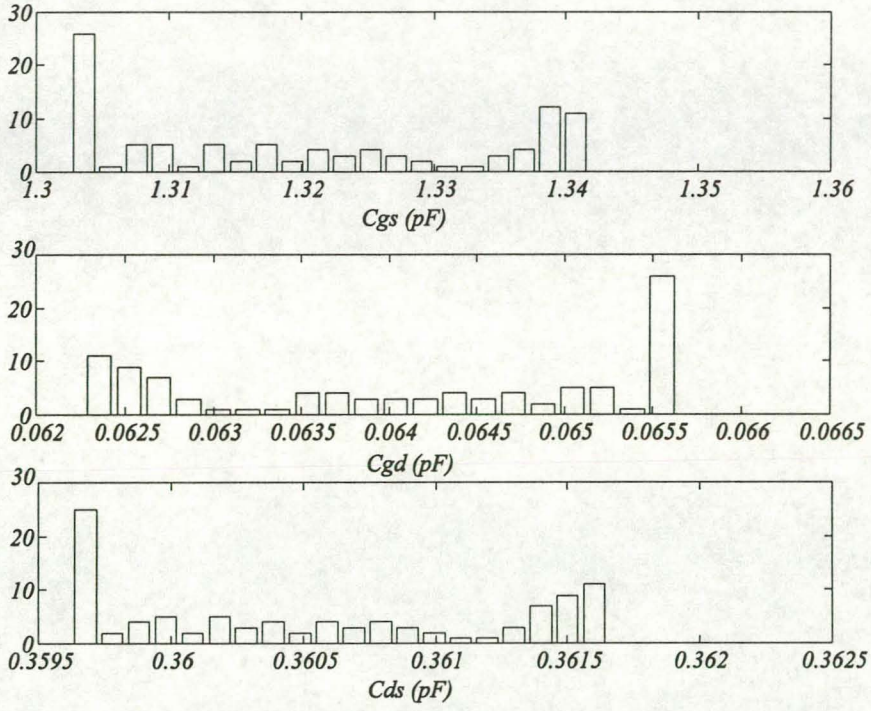


Fig. 2.7 The distribution of the FLK052XV intrinsic capacitor values. The data was generated using a 100 extractions with random starting values.

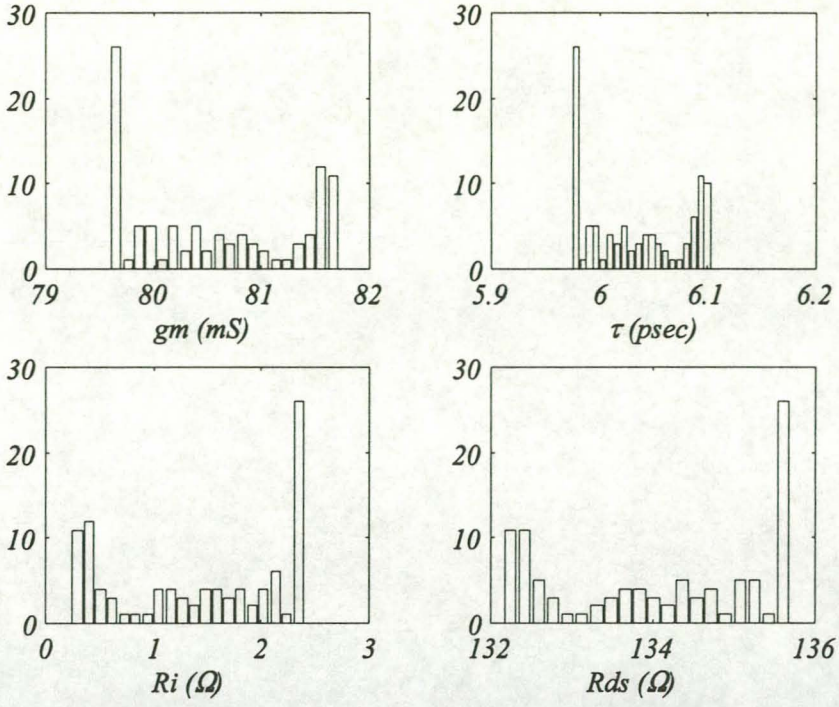


Fig. 2.8 The distribution of the intrinsic current source parameters and the channel resistances Ri and Rds for the FLK052XV transistor. The data was generated using a 100 extractions with random starting values.

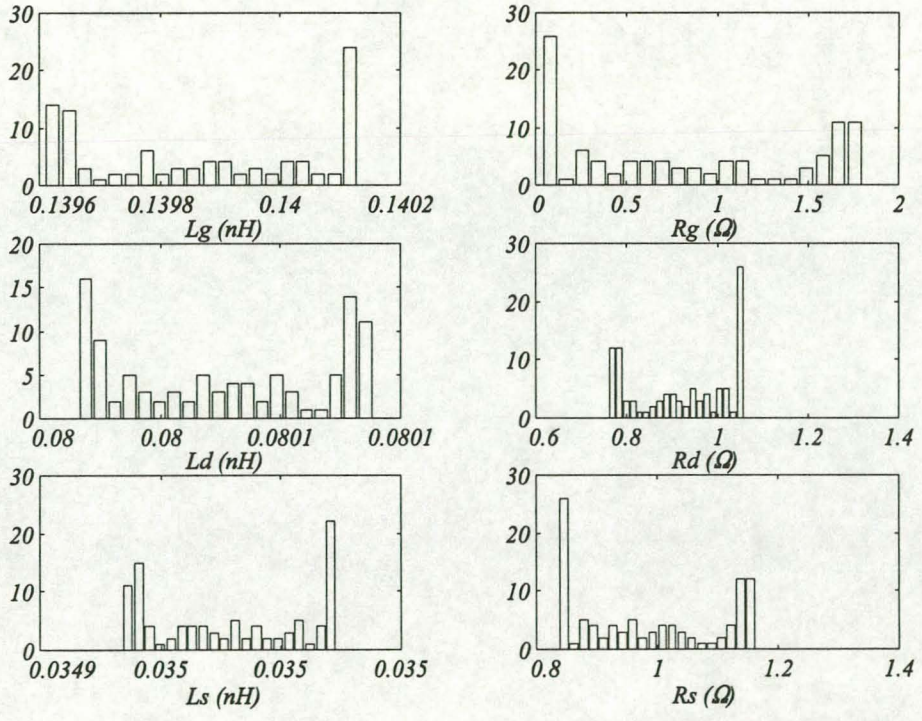


Fig. 2.9 The distribution of the extrinsic parasitic components for the FLK052XV. The data was generated using a 100 extractions with random starting values.

The influence of the size of the search space was investigated by increasing the search space range from 0.05 to 10 times the correct element values, and from 0.033 to 20 times the correct element values. As the search space was increased, the average error and the minimum and maximum values of the extracted elements increased slightly. For some examples the second search area became too large, leading to instances of instability. Instability is defined as extractions where the elements do not converge to their correct values. It can easily be detected since more than one of the dominant elements will get stuck at, or close to, one of the search boundaries. In extreme cases the final global error is also larger than the error at the starting point of the extraction.

A comparison of the results illustrates several points.

- Not all parameter extraction problems are equally accurate. This was also found by Curtice [9].
- All the elements except Rg could be found with very high accuracy and small variations in the extracted values. The FHR02X showed far larger variations in the values of the extracted elements, but the value of Rg could be found with much higher accuracy than in the other cases.
- The more dominant elements such as the intrinsic capacitors, the delay factor τ and the extrinsic inductors Lg, Ld and Ls were always found with high accuracy and very small variations in their extracted values.
- The less dominant elements such as gm and Rds were also found with high accuracy.
- The accuracy with which the parasitic resistors Rd and Rs and the intrinsic resistance Ri could be found varied from example to example. They are the least dominant elements in the model and their values influence each other.

2.5.2 Comparing Adaptive Decomposition to Kondoh's Algorithm

The first robustness test was also performed on the original algorithm proposed by Kondoh. The test was done with Rg as part of the optimisation problem and repeated with Rg fixed at its correct value. Table 2.8 contains the results. Figures 2.10 and 2.11 provides a graphical representation of the extraction results for the dominant elements Cgs and τ, and for the two non-dominant elements Ri and Rg for the FLR016XV. It can clearly be seen that τ gets caught in local minima for quite a number of starting values. Although not being the least dominant element in the model, τ is optimised last in the sequence proposed by Kondoh. This has the added effect of causing Ri to be determined incorrectly.

Table 2.8 The average and maximum percentage error in 100 extractions of the 13 element FET model from perfect data using [64], for Rg both free and fixed at its correct value.

	Rg Free				Rg Fixed			
	FLR016XV		FHR02X		FLR016XV		FHR02X	
	Avg.	Max.	Avg.	Max.	Avg.	Max.	Avg.	Max.
Cgs	0.11	1.91	3.72	63.53	3.37	71.36	0.13	0.44
Cgd	0.33	4.19	6.88	159.93	1.97	41.47	0.0027	0.42
Cds	0.24	3.24	7.03	90	0.40	6.31	1.036	4.54
gm	4.66	71.66	8.30	52.68	7.35	84.89	0.50	2.11
tau	26.13	400	10.78	393.98	29.37	400	0.16	4.80
Ri	9.19	90	124.4	400	4.44	90	1.60	7.20
Rds	0.86	13.49	5.61	138.27	1.41	15.27	0.28	1.33
Lg	0.085	4.60	2.76	51.33	11.63	269.78	0.032	0.22
Ld	0.39	5.57	12.87	90	0.95	15.57	0.93	4.17
Ls	0.36	6.73	2.34	45.33	13.48	315.86	0.11	0.45
Rg	163.84	400	134.34	400	0	0	0	0
Rd	0.71	20.55	221.48	400	0.63	25.02	19.50	84.11
Rs	2.47	19.85	35.47	400	1.85	38.43	3.99	17.47

The results summarised in table 2.8 can be misleading. The algorithm described in [64] can provide accurate extractions, but due to the fact that it is more susceptible to local minima it gets caught far from the correct solution more often. This leads to a higher average error.

A comparison of the results in table 2.8 and figures 2.10 and 2.11 with those presented in the previous section reveals the following:

- The sensitive elements of the model that are placed low down in the optimisation sequence in [64], are more likely to get caught in local minima than in the new algorithm. This will also cause the insensitive elements to be determined incorrectly.
- Fixing Rg at its correct value improves the extraction accuracy obtained for both methods, but it does not prevent the sensitive elements from getting stuck in local minima for the Kondoh algorithm.

Although the extraction accuracy obtained with [64] is still very high for most of the extractions in test one, the procedure is not as robust or accurate as the new adaptive optimisation sequence.

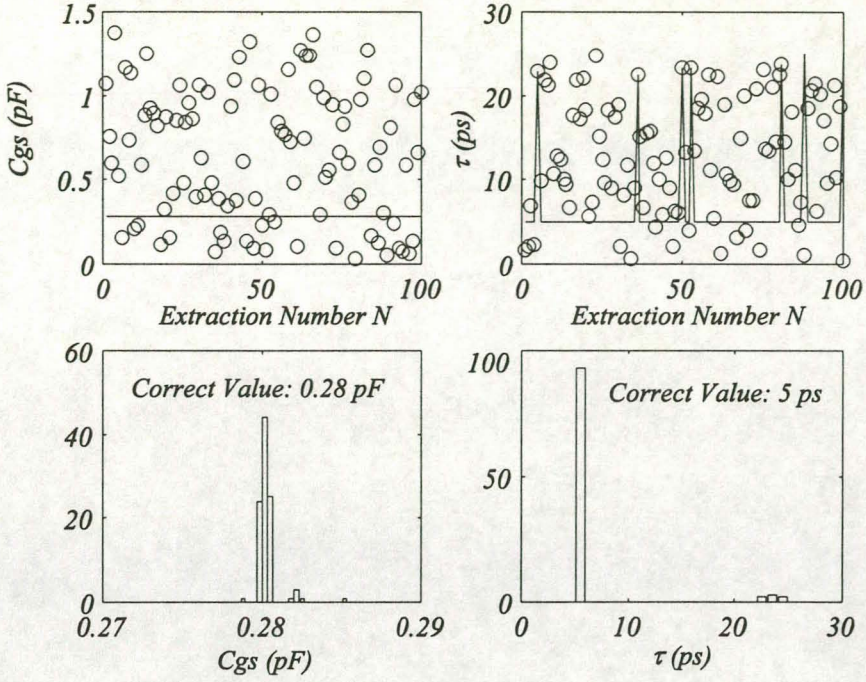


Fig. 2.10 Convergence diagram and histogram for the first robustness test using the Kondoh algorithm for Cgs and τ.

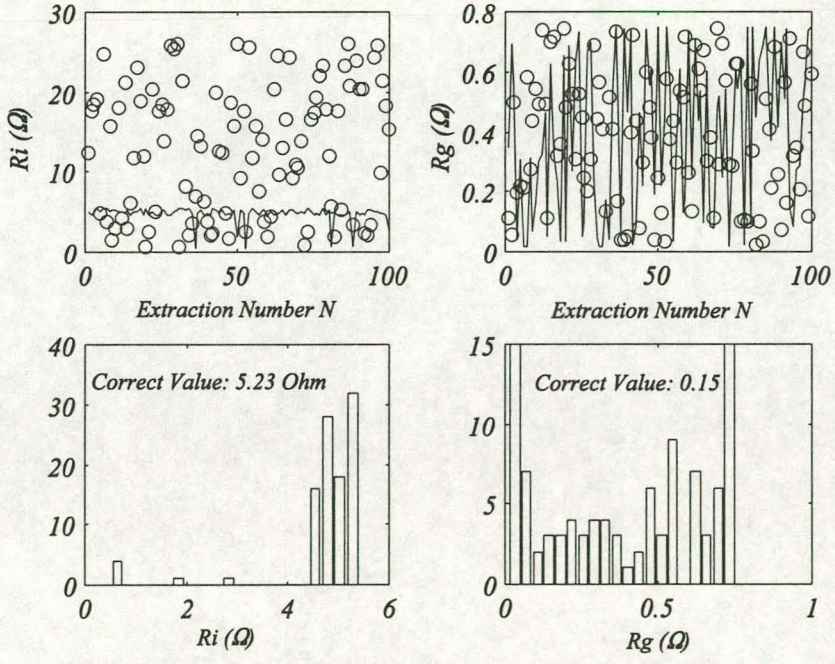


Fig. 2.11 Convergence diagram and histogram for the first robustness test using the Kondoh algorithm for Ri and Rg.

2.5.3 Extractions from Noisy Simulated Data

In order to examine the performance of the extraction algorithm when a perfect match between the model response and measurements are not possible, noise was added to the simulated data. The magnitude of the noise has a Gaussian distribution with a standard deviation equal to a chosen percentage of the measured quantity's magnitude. The phase of the noise vector is taken to have a uniform distribution between 0 and 2π radians. An example of this is graphically illustrated in figure 2.12.

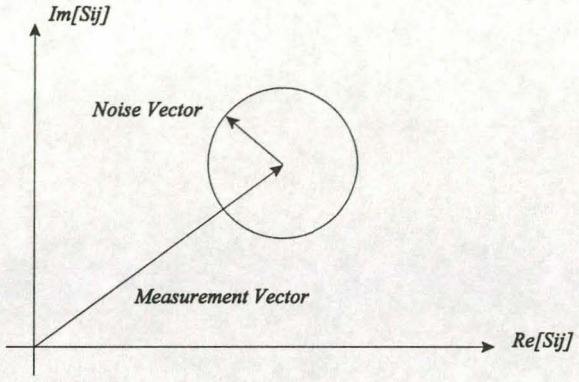


Fig. 2.12 A graphical representation of the noise vector used to corrupt the simulated measurements.

The robustness tests of the previous section were repeated, with the simulated data contaminated with a different set of noise vectors before each extraction. The tests were also repeated with the parasitic gate resistance Rg fixed at its correct value. In all the experiments the magnitude of the noise vector had standard deviation of 2% of the measurement magnitude. All the experiments were performed 500 times.

Figures 2.13 to 2.15 show histograms of the extraction results for the FHR02X. Each histogram was generated by sorting the data into 20 bins. On each histogram the normalised curve of a Gaussian distribution is superimposed. This curve has a mean and standard deviation as calculated from the extracted model element values. The results indicate, as would be expected, that the values of the dominant elements are not seriously affected by the noise, while the less sensitive elements such as Ri, Rds and the parasitic resistors Rd and Rs show large variations in their values. Figures 2.14 and 2.15 show how sensitive these elements are to measurement imperfections.

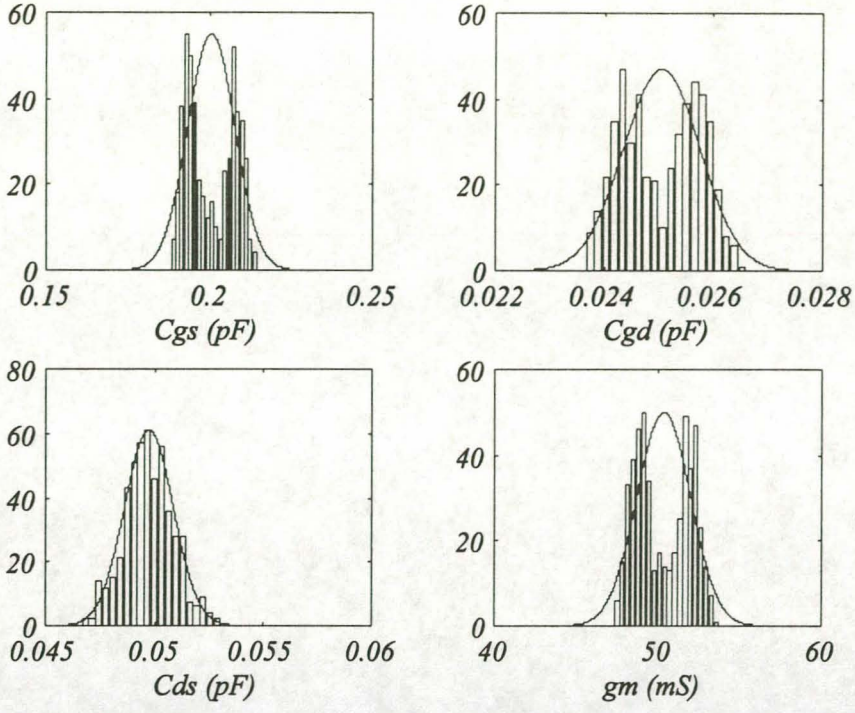


Fig. 2.13 The distribution of the intrinsic capacitors and the transconductance of the FHR02X after extractions from noisy data. Rg was not fixed at its correct value.

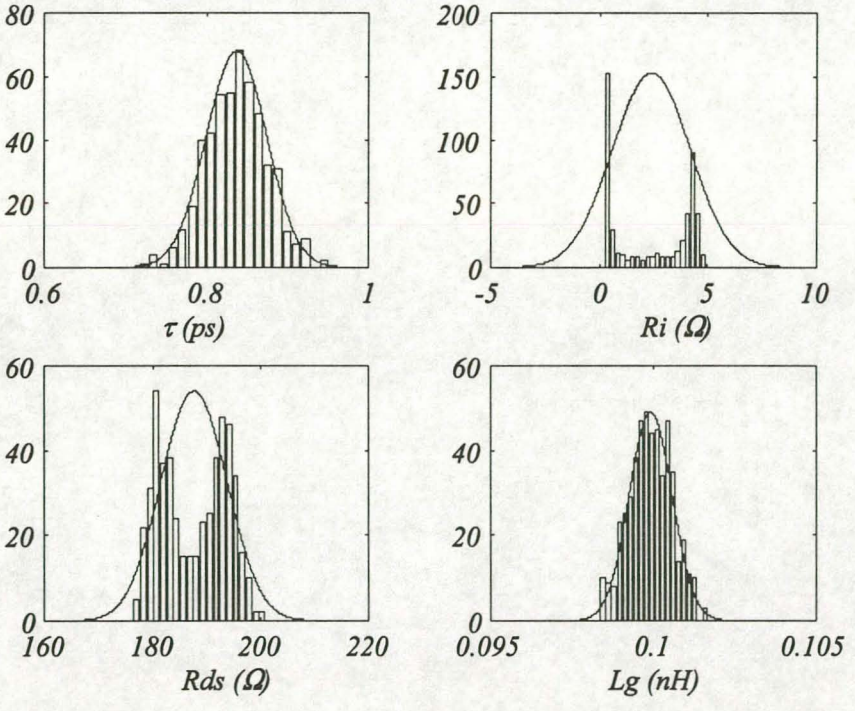


Fig. 2.14 The distribution of the channel delay τ, the channel resistors Ri and Rds and the gate inductance of the FHR02X transistor after extractions from noisy data. The value of Rg was not fixed at its correct value. Note the non-Gaussian distribution.

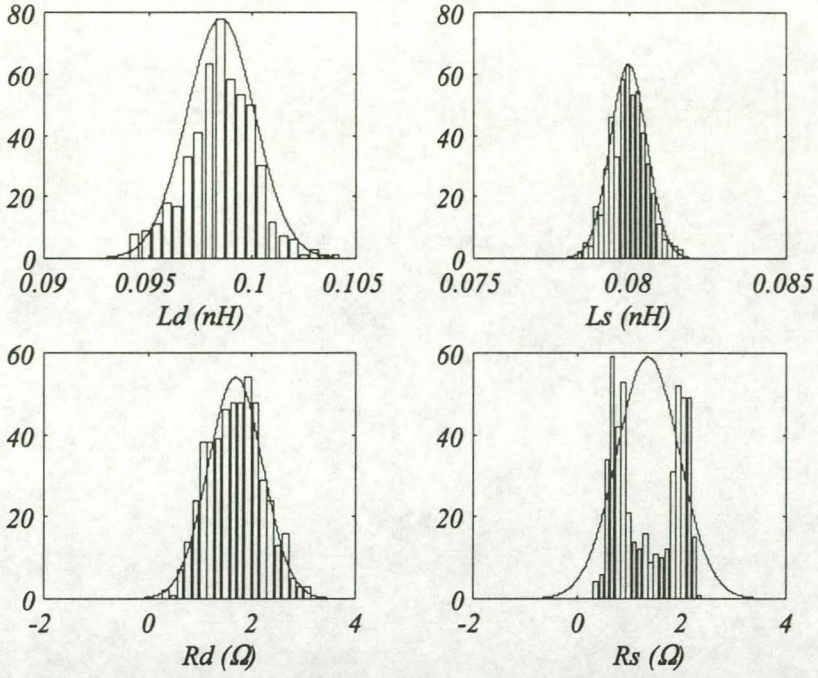


Fig. 2.15 The distribution of the extrinsic elements Ld, Ls, Rd and Rs of the FHR02X after extractions from noisy data. The value of Rg was not fixed at its correct value.

From figures 2.13 to 2.15 it can be seen that some of the element distributions appear to be Gaussian, while others cannot be described by a single normal distribution. The hypothesis that the distributions are normal was tested with the chi-square method. A significance level of α=0.05 was used. This means that if the test is passed, there is a 95% probability that the data has a Gaussian distribution [92]. Very few of the distributions shown in figures 2.13 to 2.15 passed the probability test.

Two mechanisms influence the distribution of the extracted model elements. The first is the Gaussian noise that is added to the simulated data. Since this is a random process, it is expected that its effect can be described statistically. The second mechanism is the shape of the objective function. This influences the inter element correlations and cannot be described statistically. The dominant mechanism that influences the element distribution varies from element to element and from device to device.

The experiment was also repeated with Rg fixed at its correct value. This simplifies the shape of the objective function and removes the correlation between Rg and the other elements. Figures 2.16 to 2.18 contain the new element distributions. The chi-square test confirmed that the majority of the distributions were now Gaussian. Noise is now the dominant mechanism that influences the distribution.

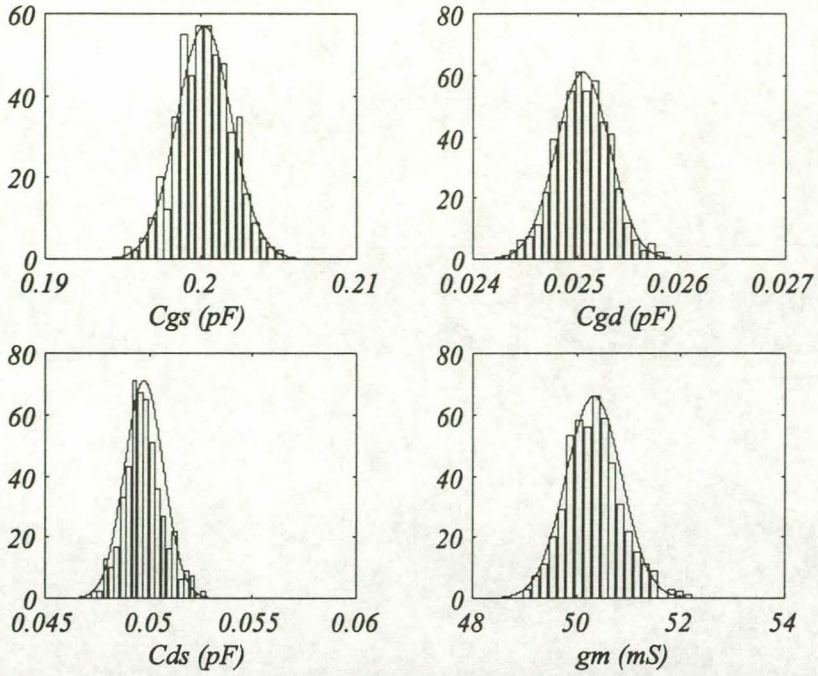


Fig. 2.16 The distribution of the intrinsic capacitors and the trans-conductance of the FHR02X transistor after extractions from noisy data. Rg was fixed at its correct value.

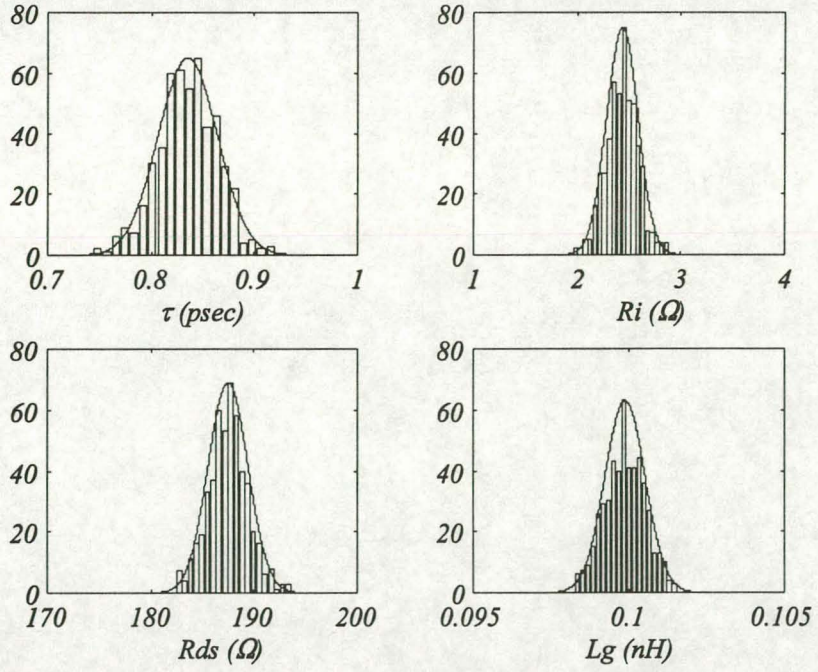


Fig. 2.17 The distribution of the channel delay τ, the channel resistors Ri and Rds and the gate inductance for the FHR02X transistor after extractions from noisy data. Rg was kept fixed at its correct value.

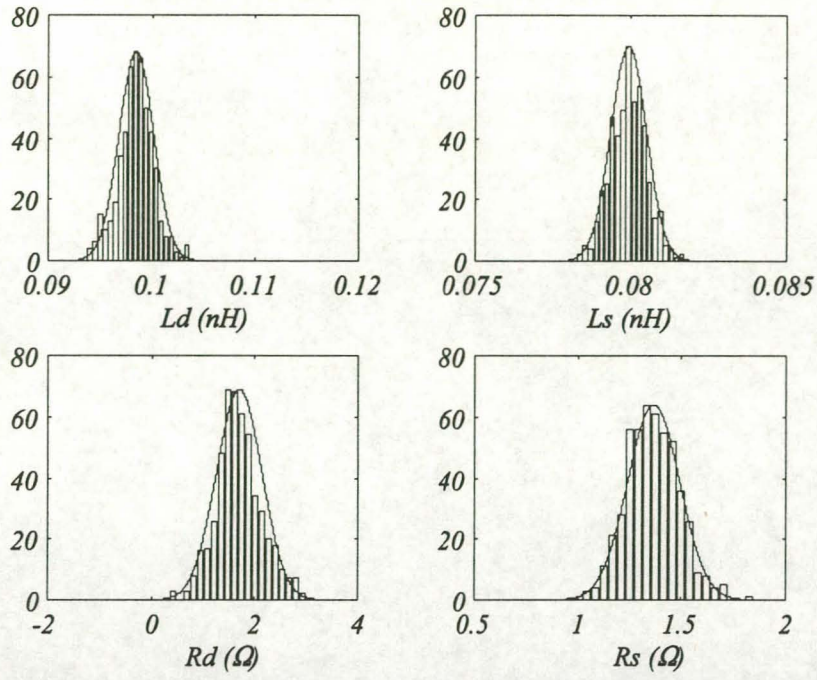


Fig. 2.18 The distribution of the drain and source inductance and parasitic resistors for the FHR02X transistor after extractions from noisy data. Rg was fixed at its correct value.

2.5.4 Extracting more Complex FET Models

It becomes necessary to extend the 13 element model when dealing with transistors that have large gate and drain networks compared to their operating frequency. Examples of such devices are packaged transistors and devices operating at mm-wave frequencies. The extensions are mostly in the form of capacitive elements that account for electromagnetic coupling between the device pads and between the pads and the ground plane. Figure 2.19 shows two different model topologies that were investigated. Figure 2.20 contains an extension for the intrinsic model, mostly found when modelling HEMT devices. The feedback resistor Rj, which can normally be ignored at lower frequencies, is included in the model to account for distributed effects under the gate. These models are all commonly found in the literature.

The extraction of more complex devices presents new problems. The 13 element model ignores the effect of geometric capacitances resulting from the extrinsic networks. It is assumed that these capacitances are absorbed in the values of the intrinsic capacitors Cgs, Cgd and Cds. Since the different capacitive effects do not have to be separated, the 13 element model is relatively well behaved and can be accurately extracted even when the upper measurement frequency is not very high. However, to accurately and consistently separate the values of the extrinsic and intrinsic capacitors, the values of the parasitic inductors and resistors, as well as the frequency

range over which measurements are taken, become important. The reason for this is simple. The extrinsic and intrinsic capacitors occur in parallel and are only separated by the other parasitic elements. Should the impedances of these separating elements become too small, the model will become non-unique.

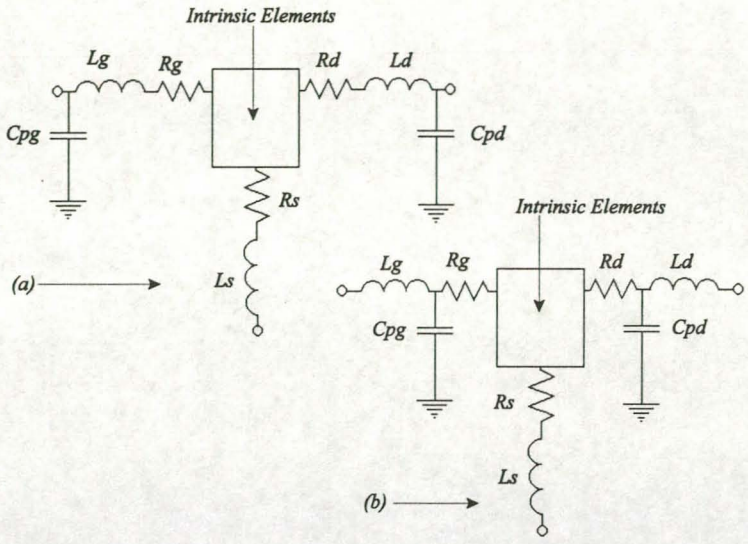


Fig. 2.19 Two different topologies that are commonly found in the literature for representing the extrinsic networks of larger FETs, and devices operating at mm-wave frequencies.

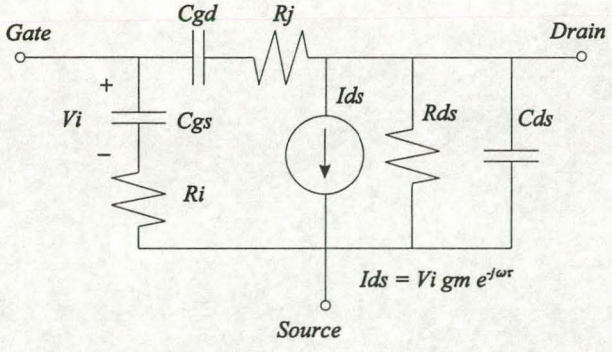


Fig. 2.20 The 8 element intrinsic network that is sometimes used to model HEMT devices.

Table 2.9 contains the model element values for two of the devices used to test the different model topologies. The values for the first transistor were taken from the literature [66] and represent an FSX03X MESFET. The second device is an extended version of the FHR02X used in the previous sections. The values of its parasitic capacitors were taken from examples found in the literature [36].

Table 2.9 The model elements for two of the devices used to study the extraction of larger models.

Parameter	FSX03X	FHR02X
Cgs (fF)	344.3	200
Cgd (fF)	32.5	25
Cds (fF)	77.4	49
gm (mS)	47.51	50
τ (psec)	2.71	0.85
Ri (Ω)	6.84	2.5
Rds (Ω)	265	188
Lg (pH)	206.8	100
Ld (pH)	202.8	100
Ls (pH)	38.4	80
Rg (Ω)	0.3	1.3
Rd (Ω)	1.3	1.3
Rs (Ω)	1.21	1.3
Cpg (fF)	38.6	20
Cpd (fF)	32.6	30

The models in figure 2.19 (a) and (b) were used to generate 40 simulated measurements between 1 and 40 GHz. A robustness test, similar to the one used for the 13 element model, was performed. A 100 random extractions were done, and the allowed search space ranged from 0.1 to 5 times the nominal element values listed in table 2.9.

Table 2.10 provides a summary of the results for the model in figure 2.19 (a). The experiments were done with Rg free and fixed at its correct value. Table 2.11 contains similar results for the FSX03X when the model in figure 2.19 (b) is used. Table 2.10 illustrates that a high level of accuracy is attainable for the expanded model shown in figure 2.19 (a). Experiments confirmed that the extraction accuracy is very much dependent on the bandwidth of the measured data, with errors in the model element values rapidly increasing as the upper measurement frequency is reduced. Fixing Rg at its correct value still results in a decrease of the errors made in determining the values of the other model elements. However, the FHR02X example shows that this trend is not assured for all the model elements. The higher dimensions of the problem are now becoming the dominant factor in the success of the parameter extraction.

Table 2.10 The results of the robustness test for the model shown in figure 2.19 (a). All values express the errors made in determining each model element as a percentage of the correct element value.

	FSX03X						FHR02X					
	Rg is free			Rg is fixed			Rg is free			Rg is fixed		
	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Lg	0.1	-0.26	0.06	0	0	0	0.38	-1.38	1.87	0.7	-1.6	2.05
Rg	154.9	-90	400	—	—	—	43.43	-90	101.3	—	—	—
Cgs	1.25	-0.71	3.21	0.04	0	0.21	1.62	-4.38	3.9	0.15	-1.98	2.33
Ri	10.33	-26.55	5.93	0.24	-0.99	0	39.17	-90	86.76	0.37	-8.11	11.81
Cgd	0.84	-2.46	0.65	0.08	0	0.26	1.4	-3.55	3.45	0	-1.15	1.53
gm	1.27	-0.63	3.27	0.19	0	0.83	0.9	-6.11	1.82	1.01	-4	6.11
τ	0.5	-0.6	1.6	0.12	-0.49	0	0.76	-2.66	3.88	0.64	-3.95	2.65
Rds	0.83	-2.67	0.92	0.18	0	0.65	1.03	-2.29	5.79	0.8	-4.15	3.6
Cds	0.89	-0.14	2.47	0.59	0	2.47	0.77	-8.47	11.34	2.41	-8.08	15.31
Rd	2.57	-31.69	32.78	8.83	-0.02	36.16	32.56	-90	169.3	33.92	-90	166.1
Ld	0.39	-1.52	0.1	0.38	-1.56	0	1.93	-14.2	7.63	3.31	-17.1	9.2
Rs	20.06	-10.76	50.7	2.06	-0.03	8.86	16.34	-75.47	35.49	11.89	-42.5	57.72
Ls	0.34	-0.86	0.17	0.11	-0.53	0	0.56	-3.41	2.61	0.79	-4.3	3.22
Cpg	0.08	-0.22	0.02	0.04	-0.18	0	1.27	-4.94	6.4	2.79	-6.29	7.68
Cpd	1.11	-3.89	0.01	0.97	-4.02	0	2.04	-13.98	7.53	3.1	-17.09	8.91

Table 2.11 shows that the model of figure 2.19 (b) failed to allow consistent extractions of the model elements. The values of the intrinsic and extrinsic capacitors could not be separated, and errors made in their determination negatively influences the other model elements. Fixing Rg at its correct value provided no improvements. It is concluded that this topology is not suited for the accurate experimental characterisation of devices.

Table 2.11 The results of the robustness for the model topology shown in figure 2.19 (b). All values express the errors made in determining an element as a percentage of the correct element value. Results are only shown for the FSX03X device.

	All model elements are free to vary			Rg is fixed at its correct value		
	FSX03X			FSX03X		
	Avg	Min	Max	Avg	Min	Max
Cgs	26.911	-37.906	11.972	27.18	-39.432	12.437
Cgd	29.825	-60.292	5.706	31.249	-60.207	4.011
Cds	5.978	-90	53.536	7.937	-90	54.979
gm	1.826	-5.822	2.193	1.693	-5.561	1.557
τ	63.2	-90	13.039	64.274	-90	12.629
Ri	139.01	-90	233.428	141.71	-85.021	261.716
Rds	7.713	-24.016	10.957	8.415	-23.943	12
Lg	9.399	-17.868	16.612	9.586	-17.569	16.58
Ld	0.369	-13.154	11.375	0.704	-14.397	11.94
Ls	49.61	-21.901	107.474	51.451	-22.192	104.096
Rg	9.934	-90	400	—	—	—
Rd	62.918	-90	400	69.905	-90	400
Rs	80.92	-90	39.772	80.962	-90	8.393
Cpg	266.8	-90	400	271.87	-90	400
Cpd	60.971	-90	209.31	68.093	-90	210.099

The 15 element FSX03X model was modified in order to include a feedback resistance Rj of 150 Ω. A robustness test, similar to those described in the previous paragraphs were performed for the two cases where Rg is free to vary and when Rg is fixed at its correct value. The experiments were also repeated using noisy simulated data. The noise that was added had the same specifications as that used in section 2.5.3. The starting values for the elements were chosen in a range of ±99% of the nominal element values when noisy data was used. The results for the different experiments are listed in table 2.12.

Table 2.12 The results of the robustness test for the 16 element HEMT model. All values express the errors made in determining an element as a percentage of the correct element value.

	Perfect Data						Noisy Simulated Data					
	Rg Free			Rg Fixed			Rg Free			Rg Fixed		
	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg	Min	Max
Cgs	0.08	-5.289	1.299	0.112	-5.236	0.254	0.908	-5.857	2.243	0.727	-5.773	1.499
Cgd	0.514	-1.348	2.515	0.492	-0.633	2.426	0.5	-2.389	3.577	0.464	-0.938	3.214
Cds	2.096	-9.352	6.906	1.943	-6.499	6.351	1.046	-10.84	7.764	0.115	-8.725	6.348
gm	0.627	-5.403	1.974	0.535	-5.219	2.037	0.765	-6.427	2.908	0.522	-5.993	2.293
τ	0.506	-1.92	3.209	0.504	-1.575	2.043	0.547	-3.22	4.317	0.02	-3.02	4.043
Ri	1.426	-20.1	14.2	0.309	-2	9.884	3.443	-22.11	13.44	1.37	-3.71	12.34
Rj	1.194	-39.67	2.599	0.764	-31.2	1.944	8.249	-39.43	6.694	4.288	-32.14	6.529
Rds	0.657	-0.78	5.167	0.618	-0.74	5.059	1.282	-3.003	6.69	0.865	-3.356	6.483
Lg	0.09	-1.029	0.314	0.08	-0.961	0.319	0.133	-2.203	0.959	0.08	-1.626	1.125
Ld	1.779	-4.767	2.128	1.573	-4.546	2.044	1.025	-5.883	2.661	0.859	-5.823	2.804
Ls	0.545	-2.023	9.115	0.468	-2.125	8.701	1.493	-2.977	9.791	0.92	-2.753	9.315
Rg	21.33	-90	400	—	—	—	109.4	-90	400	—	—	—
Rd	49.12	-69.48	149.1	44.97	-57.06	136.6	5.496	-90	170.9	16.03	-89.22	143.2
Rs	3.615	-90	16.95	2.589	-90	17.03	14.51	-90	29.69	11.02	-90	23.4
Cpg	0.151	-0.344	0.08	0.135	-0.326	0.08	0.03	-3.725	4.338	0.217	-5.091	3.672
Cpd	3.219	-8.292	3.866	2.835	-8.106	3.607	2.098	-12.04	6.224	4.288	-32.14	6.529

Table 2.12 shows that the algorithm still provides surprisingly accurate results when models with a very large number of elements are extracted. The experiments with noisy data were performed to see if the extractions become overly sensitive to deviations in the data when such large models are used. A comparison of the results for the 13, 15 and 16 element models indicate that there is an increase in modelling errors with the increase in the number of model elements. This is to be expected since the values of the extracted elements are also a function of the errors made in the determination of the other elements. The principal components analysis shows that the feedback resistance Rj is a very insensitive element (like Ri), but due to its larger value, it can be determined with reasonable accuracy.

One important difference between the more complex models and the 13 element model is the influence that Rg has on the extraction accuracy. A comparison of the results show that the accuracy improvement, resulting from fixing Rg at its correct value, decreases for the larger models. The distribution of the model elements extracted from noisy data also reflects the influence of the increased model complexity. Figure 2.21 shows that the distributions are non-Gaussian, regardless of whether Rg is fixed at its correct value or not. The more complex shape of the objective function now has a far more dominant effect on the distribution of the model elements than the noise that is added to the simulated s-parameters.

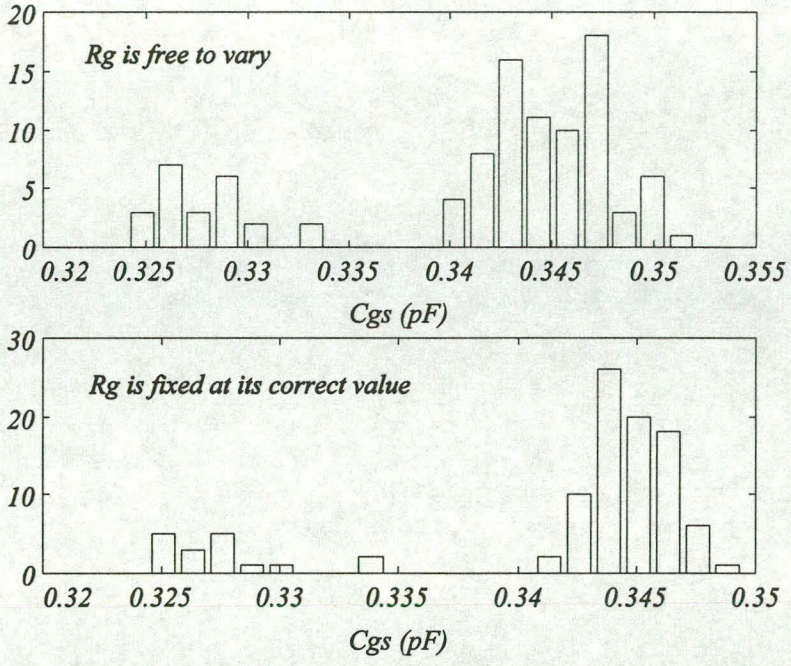


Fig. 2.21 The distribution of the extracted Cgs values from noisy simulated data for the FSX03X device. A 16 element model was used and the experiment was repeated with Rg free and fixed at its correct value.

The decomposition-based method was also tested on cold FET models. These models are commonly used to determine the extrinsic elements of a FET. When the FET is biased with V_{GS} smaller than the pinch-off voltage and V_{DS} = 0 Volt, the intrinsic model simplifies to three capacitors (Cgs, Cgd and Cds). The cold FET model contains fewer elements, but also presents 25% less measured data since the FET now acts as a passive circuit. This causes the subfunction/model element associations to become non-unique, with the assignments being very much dependent on the specific topology used in the model equivalent circuit. The decrease in the available subfunctions, and the uneven distribution of the model elements among the available functions, cause the algorithm to fail when applied to cold FET data. The algorithm performs best when used with non-reciprocal device measurements.

2.6 Model Extractions from Measured Data

Two different devices, a GaAs MESFET and a GaAs pseudomorphic HEMT (pHEMT), were used to verify the single bias extraction algorithm on measured data.

2.6.1 GaAs MESFET

The GaAs MESFET is a 0.25µm × 300µm low noise device made by Texas Instruments. The s-parameter data for the MESFET was measured over a wide range of bias points. Figure 2.22 shows the I_{DS}-V_{DS} curves for the transistor. The V_{GS} bias voltage is varied from -1.6 Volt to 0.4 Volt in steps of 0.2 Volt. Each marker in figure 2.22 represents a bias point at which s-parameters were measured. All the s-parameter measurements ranged from 45 MHz to 26.5 GHz. The measurements were made at 51 frequency points. The s-parameter measurements were performed using an HP8510 network analyser [93].

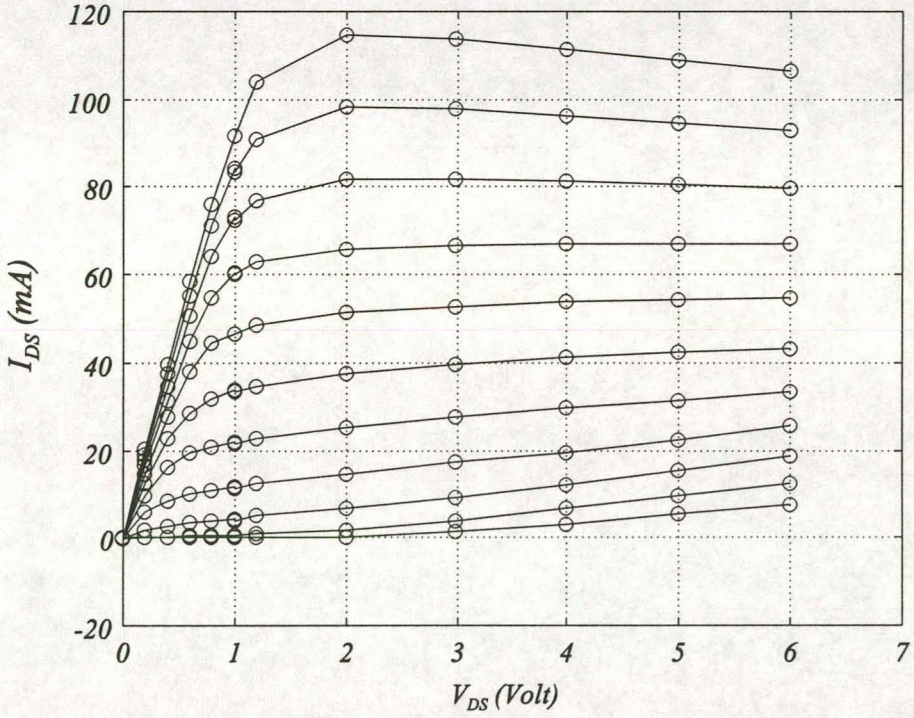


Figure 2.22 The I_{DS}-V_{DS} curves for the 0.25µm × 300µm low noise GaAs MESFET used in the extractions. Each marker represents a bias point at which s-parameters were measured. The V_{GS} voltage varied from -1.6 Volt to 0.4 Volt in steps of 0.2 Volt.

Starting values for the extraction were determined by choosing a bias point in the middle of the I_{DS}-V_{DS} plane (V_{GS} = -0.8 Volt and V_{DS} = 3 Volt) and extremely wide optimisation boundaries. A small number of random extractions were then performed, the results of which provided more

accurate nominal element values. The starting points were now chosen in a range of -99.99% to 400% around these nominal element values. The optimisation space was made large in order to ensure that it contains the extraction solutions for a wide range of bias points.

Experiments were performed using both 13 and 15 element models. The 15 element model failed to provide better modelling accuracies than the 13 element model and was unable to converge to a unique solution. The intrinsic elements extracted with the 15 element model were also not physically believable, with gm converging to very high values and τ to very low values. Further MESFET modelling was subsequently performed using only the 13 element model.

A robustness test, consisting of 100 extractions started from random starting values uniformly distributed in the optimisation space, was done for a range of bias voltages. The extractions included the bias points from V_{GS} = -1.0 Volt to V_{GS} = 0.0 Volt and V_{DS} = 1.0 Volt to 6.0 Volt. Figures 2.23 to 2.29 present the result of the extractions. Each figure contains the element values for the best extraction, was well as the average element values. It is clear from the presentation that for most of the elements, the best and average element values are very close to each other. The largest differences are seen for the channel charging resistance Ri and the parasitic extrinsic resistors Rg, Rs and Rd. These elements are extremely sensitive to measurement errors and deviations in the other model elements.

Figures 2.23 to 2.26 show that the average element values provide far smoother curves for the intrinsic element bias dependencies, even for the element Ri. The extrinsic elements are supposed to be bias independent, but from the results in figure 2.26 to 2.29 it can be seen that there is an unacceptably large variation in their values. Note that the extrinsic resistance Rd exhibits a very large bias dependence, and for high V_{DS} voltages, the element will converge to a negative value if it is allowed to.

For the majority of the bias points the distributions of the extracted model elements are small. Figure 2.30 shows the distribution of Cgs, τ, Ld and Ri for the bias point V_{GS} = -0.8 Volt and V_{DS} = 3.0 Volt. It was found that there is a large increase in the extraction uncertainty for bias points in the region of the knee-voltage of the I_{DS}-V_{DS} plane. Figure 2.31 contains the distributions of Cgs, τ, Ld and Ri for the bias point V_{GS} = 0.0 Volt and V_{DS} = 1.0 Volt.

Comparisons were also made between the measured and modelled s-parameters to determine the accuracy of the model-based s-parameter predictions. Figures 2.32 and 2.33 show the modelled and measured magnitude and phase of the four s-parameters for the bias point V_{GS} = -0.8 Volt and V_{DS} = 3.0 Volt. The predicted s-parameters in figures 2.32 and 2.33 were generated using

the average element values. The differences between s-parameters generated using the average element values and the best element values are negligible. The modelled s-parameters fit very well onto the measured data, with the largest deviations occurring in the magnitude of the s-parameters. Tests at other bias points showed similar modelling accuracies, with the largest errors occurring at high V_{GS} and V_{DS} voltages in the magnitude of S_{22}.

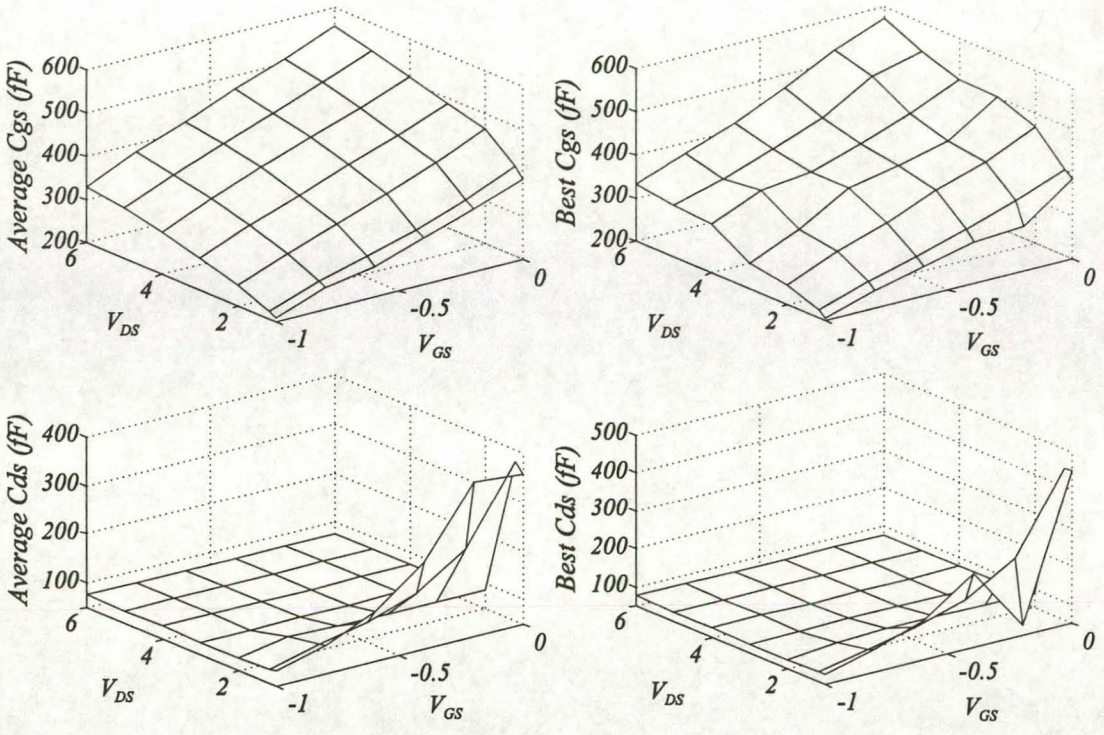


Fig. 2.23 The bias dependence of the MESFET intrinsic capacitors Cgs and Cgd.

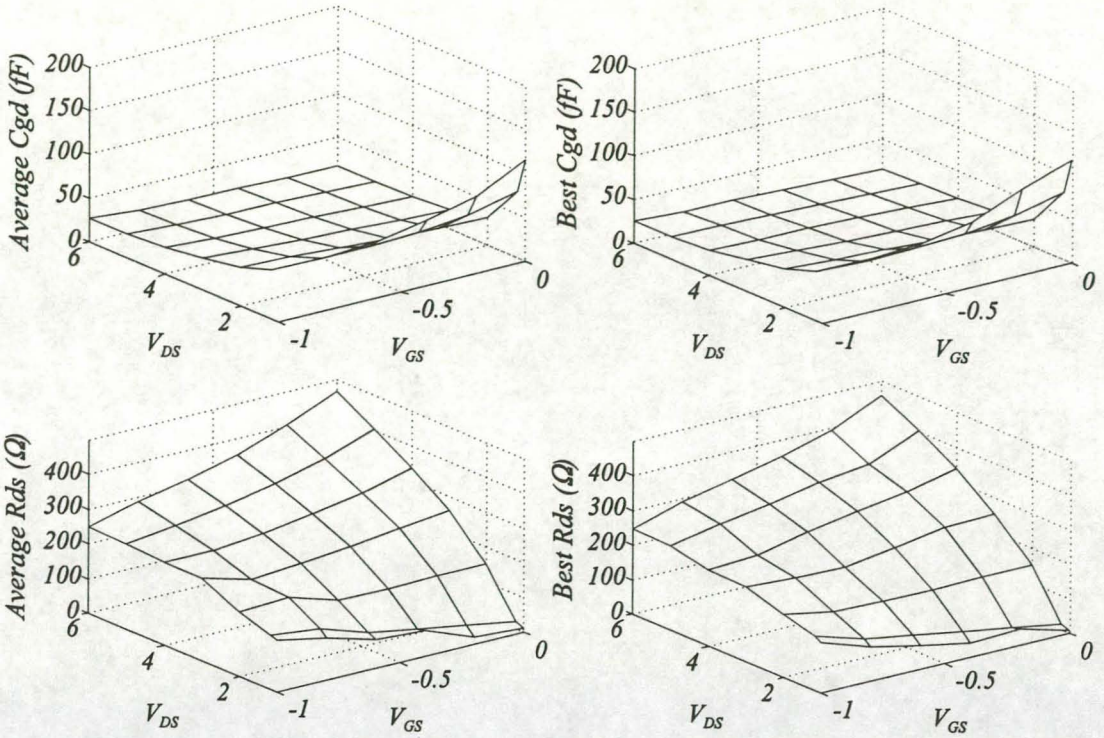


Fig. 2.24 The bias dependence of the MESFET intrinsic elements Cgd and Rds.

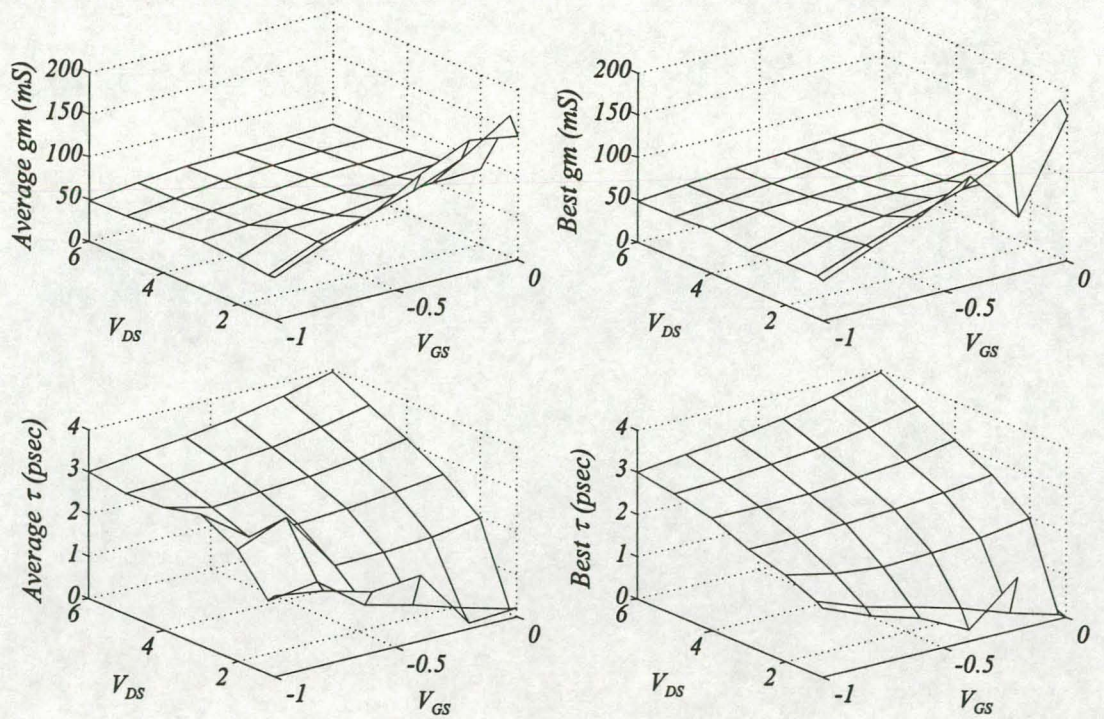


Fig. 2.25 The bias dependence of the MESFET transconductance gm and channel delay τ.

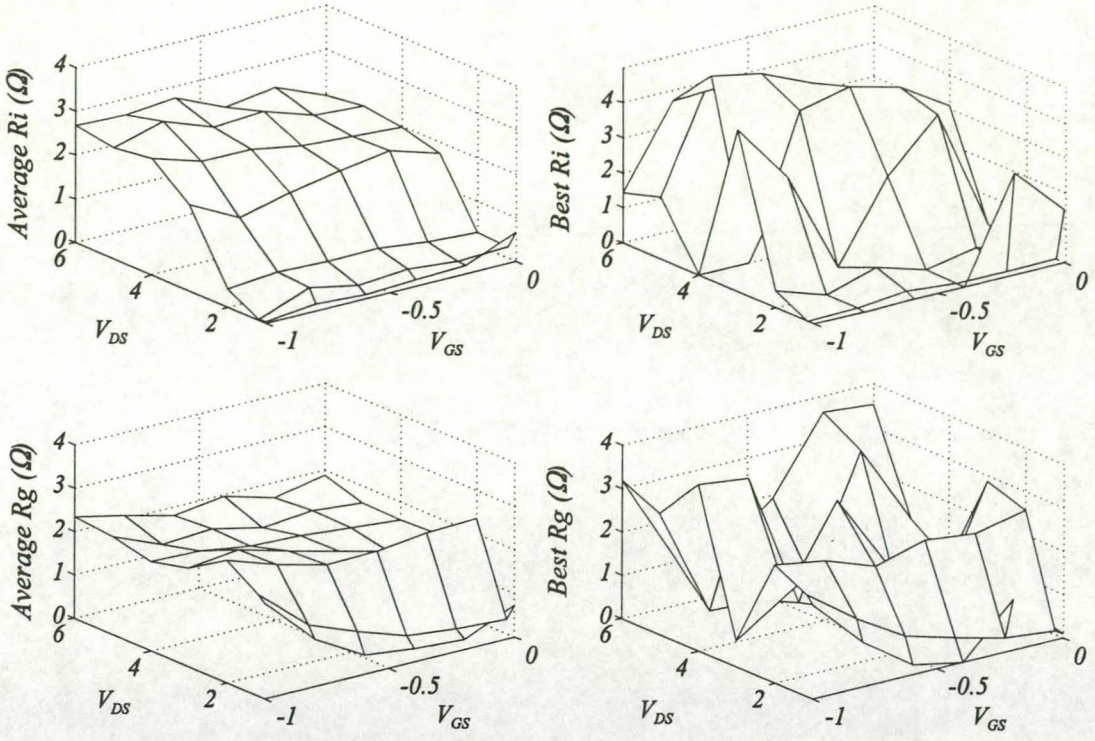


Fig. 2.26 The bias dependence of the MESFET intrinsic channel resistance Ri and the gate parasitic resistance Rg.

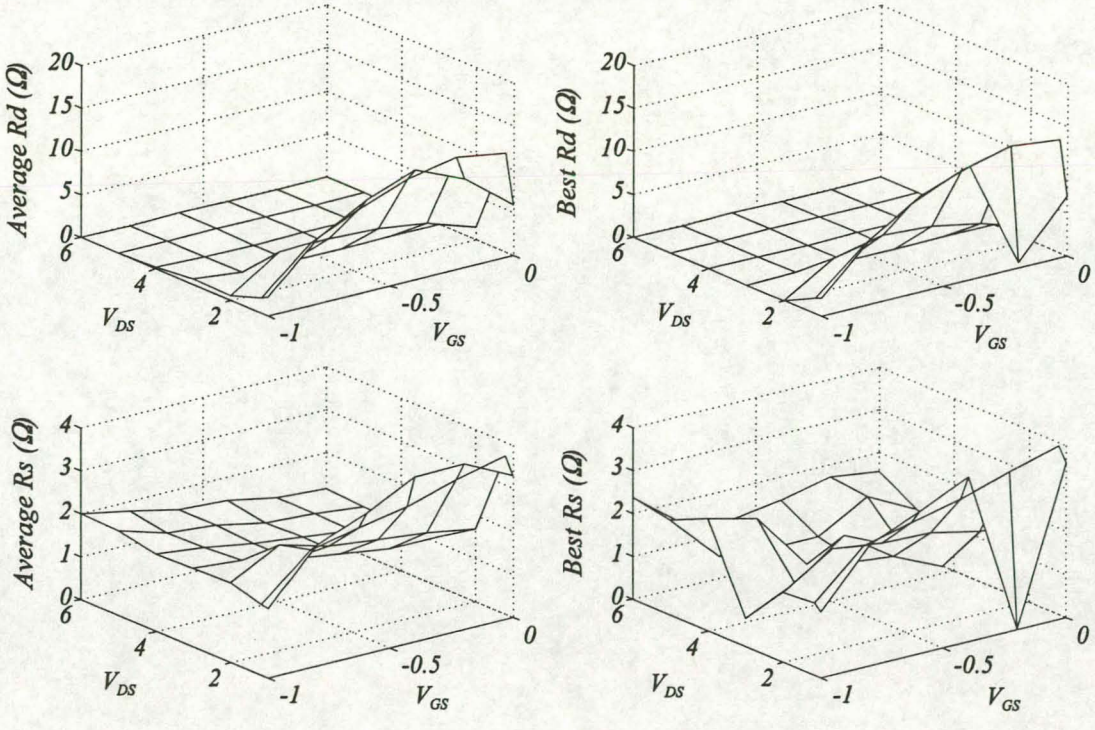


Fig. 2.27 The variation of the MESFET parasitic resistances Rd and Rs with bias.

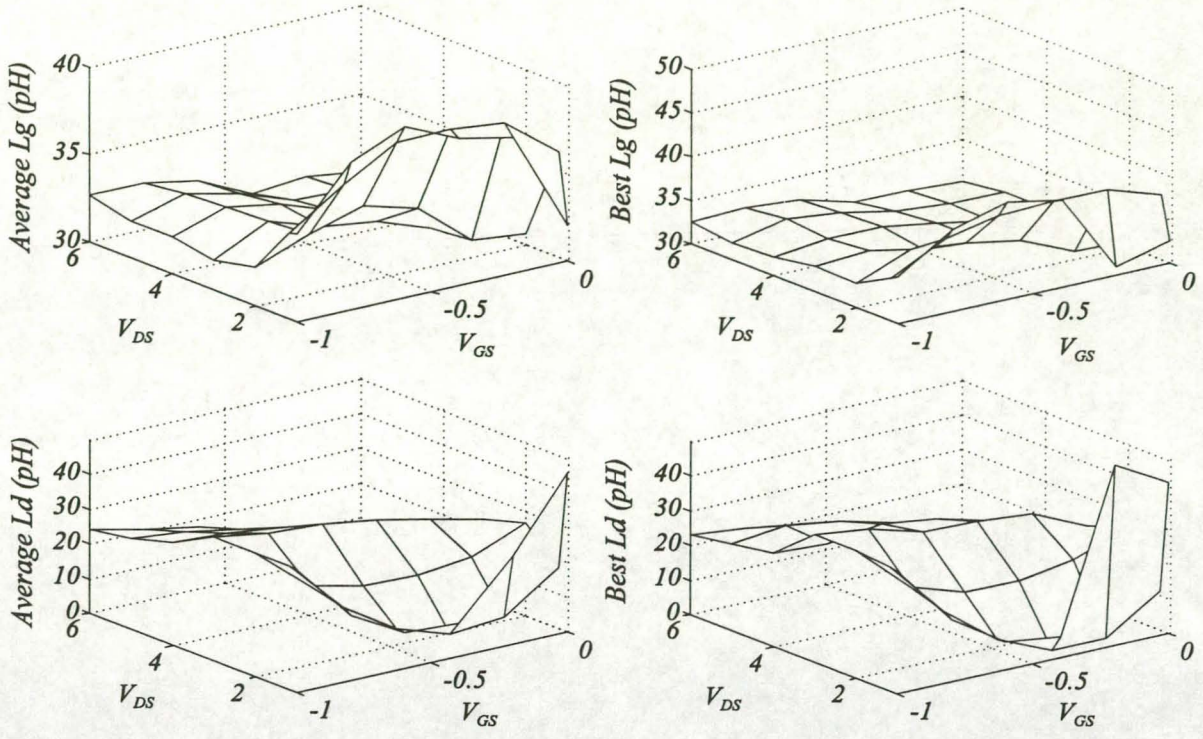


Fig. 2.28 The variation of the MESFET parasitic inductors Lg and Ld with bias.

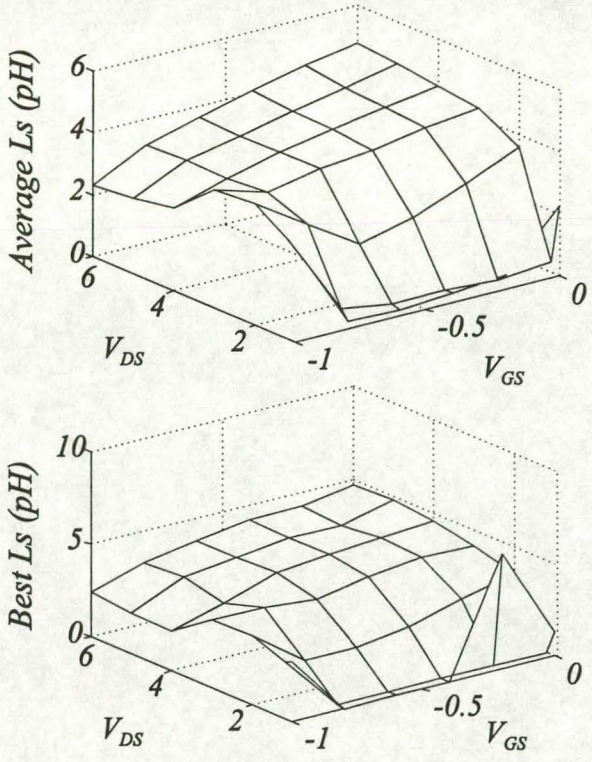


Fig. 2.29 The variation of the MESFET source inductance Ls with bias.

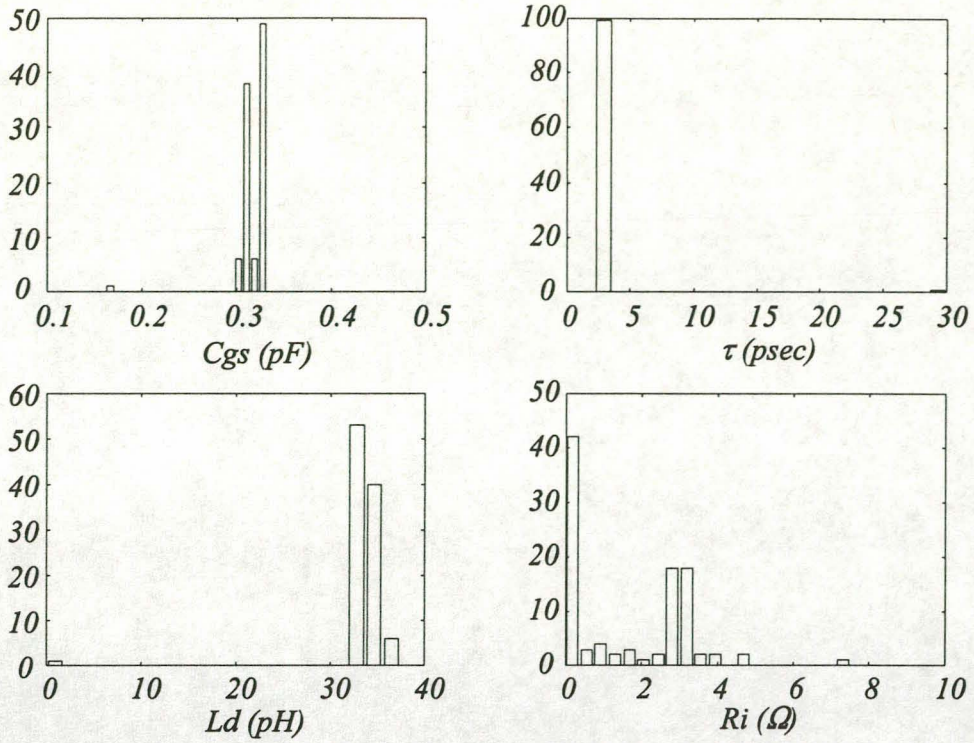


Fig. 2.30 The distribution of the MESFET elements Cgs, τ, Ld and Ri after a robustness test at bias point V_{GS} = -0.8 Volt and V_{DS} = 3.0 Volt.

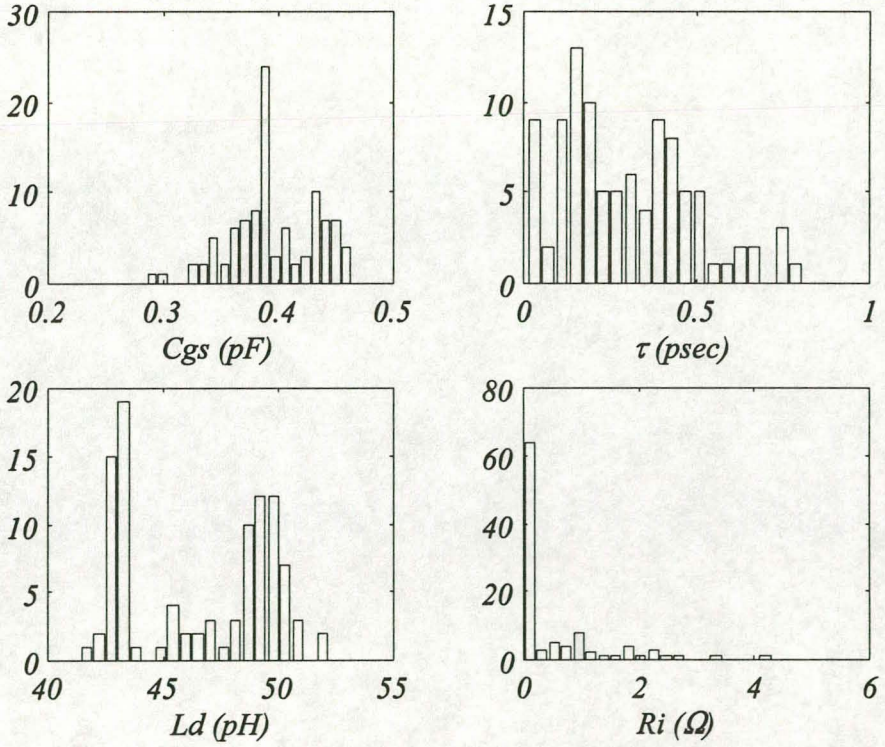


Fig. 2.31 The distribution of the MESFET elements Cgs, τ, Ld and Ri after a robustness test at the bias point V_{GS} = 0.0 Volt and V_{DS} = 1.0 Volt.

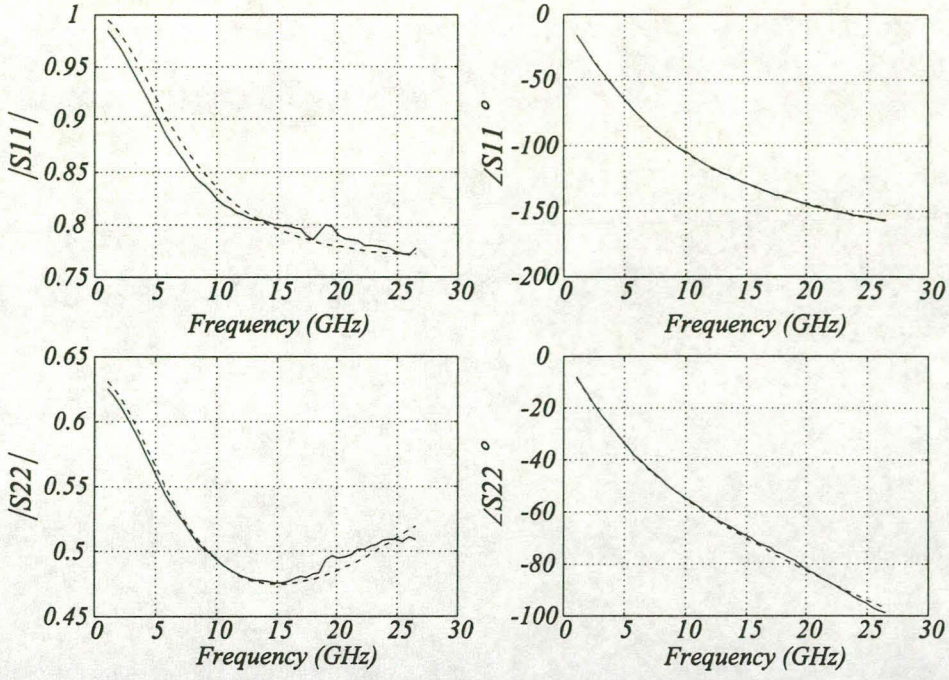


Fig. 2.32 A comparison of the modelled and measured MESFET s-parameters at bias point V_{GS} = -0.8 Volt and V_{DS} = 3.0 Volt. (___ measurement, ---- model)

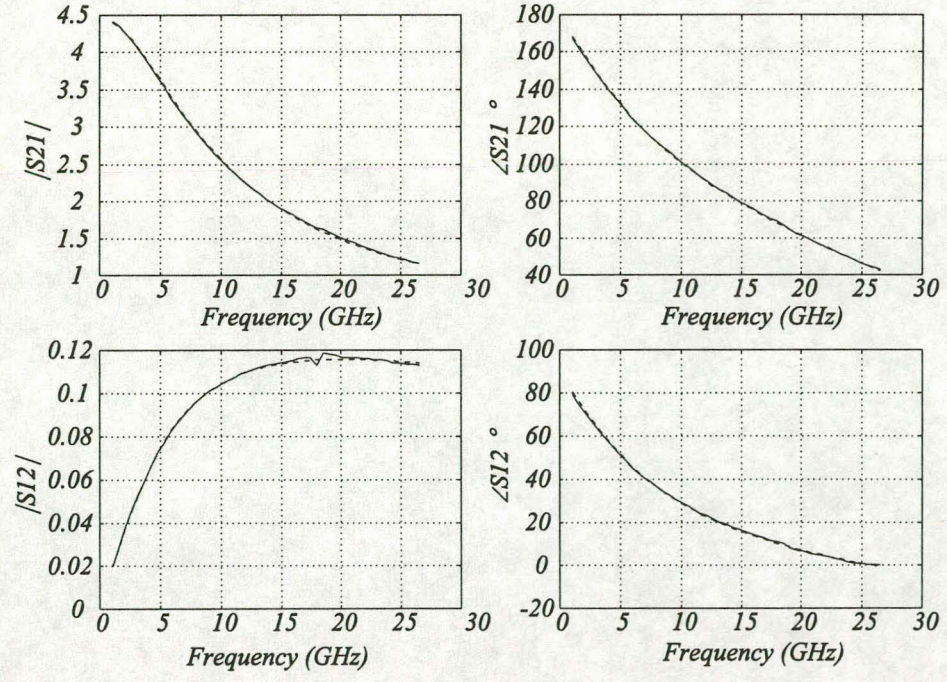


Fig. 2.33 A comparison of the modelled and measured MESFET s-parameters at bias point V_{GS} = -0.8 Volt and V_{DS} = 3.0 Volt. (___ measurement, ---- model)

2.6.2 GaAs pHEMT

The second device used for testing the extraction algorithms was a 0.2µm × 100µm GaAs pseudomorphic HEMT (pHEMT) [94]. Figure 2.34 shows the I_{DS}-V_{DS} curves, with each marker representing a point at which s-parameters were measured. The V_{GS} voltage was varied from -0.5 Volt to 0.7 Volt in steps of 0.05 Volt, and V_{DS} was varied from 0 Volt to 2.5 Volt, also in steps of 0.05 Volt. The s-parameter measurements were made at 51 frequency points, ranging from 1 GHz to 25 GHz. The measurements were performed using an HP8510 network analyser and an on-wafer measurement set up [18,95].

The same procedure as for the MESFET was employed to generate nominal element values around which a large search space could be defined. The reference bias point used for determining the nominal element values was V_{GS} = 0.0 Volt and V_{DS} = 1.0 Volt. Both a 13 and a 15 element model were used. The 15 element model once again failed to converge to a unique solution.

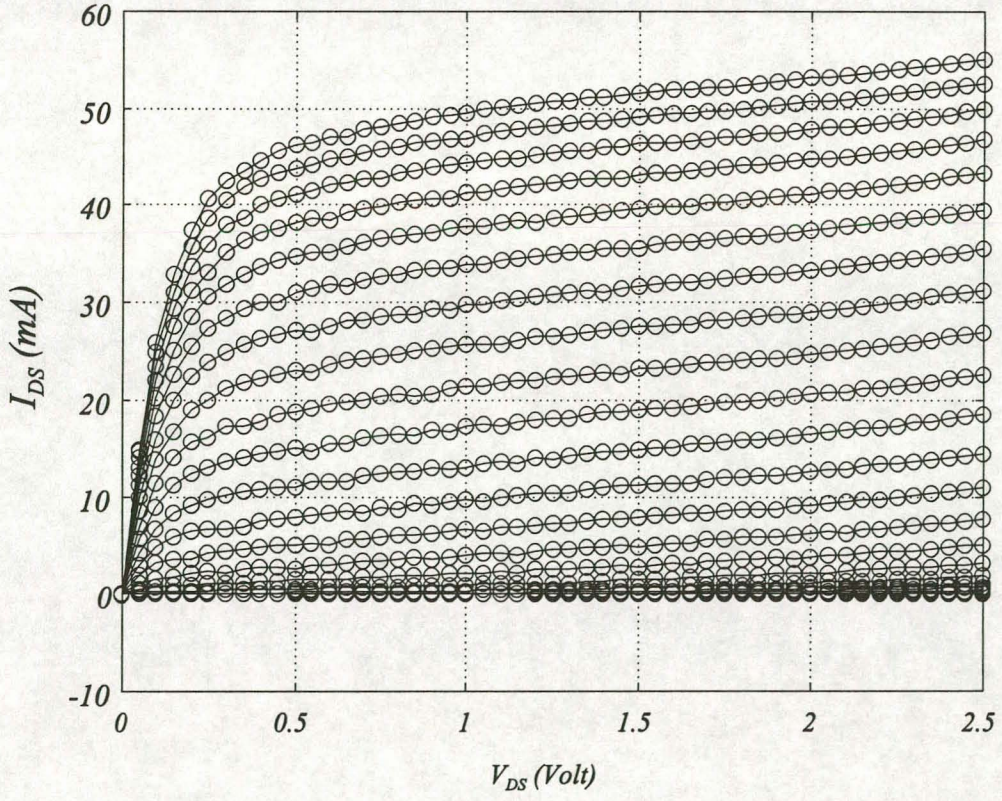


Figure 2.34 The I_{DS}-V_{DS} curves for the 0.2µm × 100µm GaAs pHEMT. Each marker represents a bias point at which s-parameters were measured. The drain and gate voltages were changed in 0.05 Volt steps.

A robustness test was done using the following bias voltages.

V_{GS} = -0.05, 0.00, 0.10, 0.20, 0.30
V_{DS} = 0.40, 0.50, 1.00, 1.50, 2.00, 2.50

Figures 2.35 to 2.41 show the best and average element values obtained in the test, while figure 2.42 and 2.43 show the distribution of Cgs, τ, Ld and Ri for two different bias points. It is again noted that there is a large increase in the extraction uncertainty for bias points near the knee-voltage.

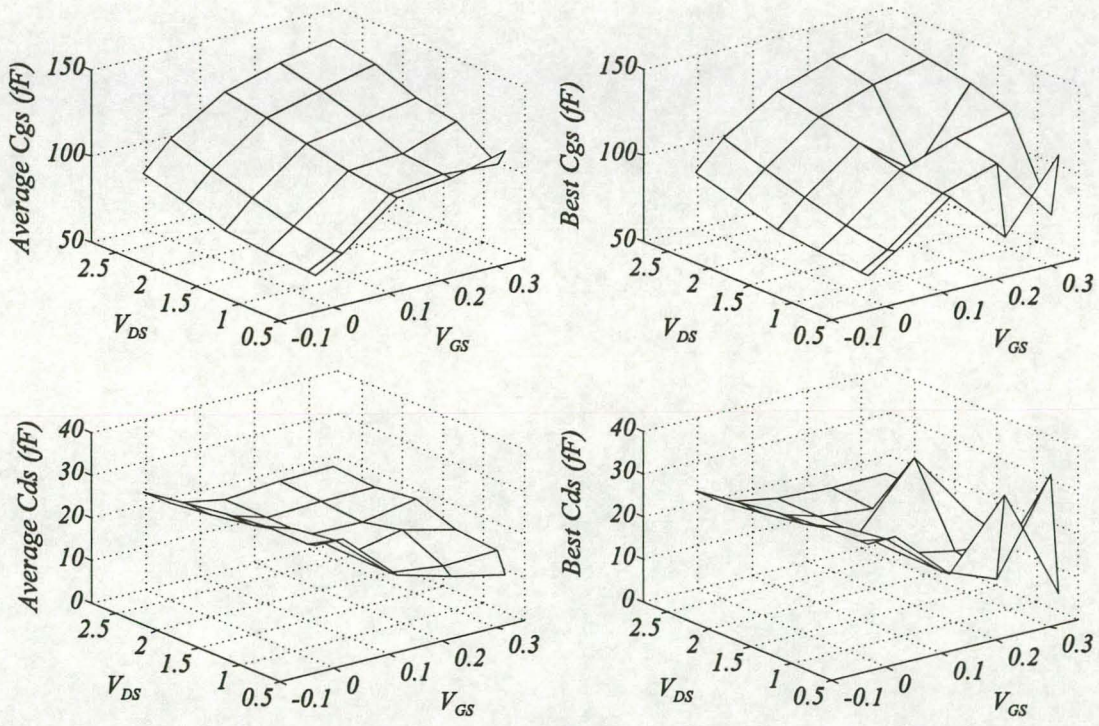


Fig. 2.35 The bias dependence of the pHEMT intrinsic elements Cgs and Cds.

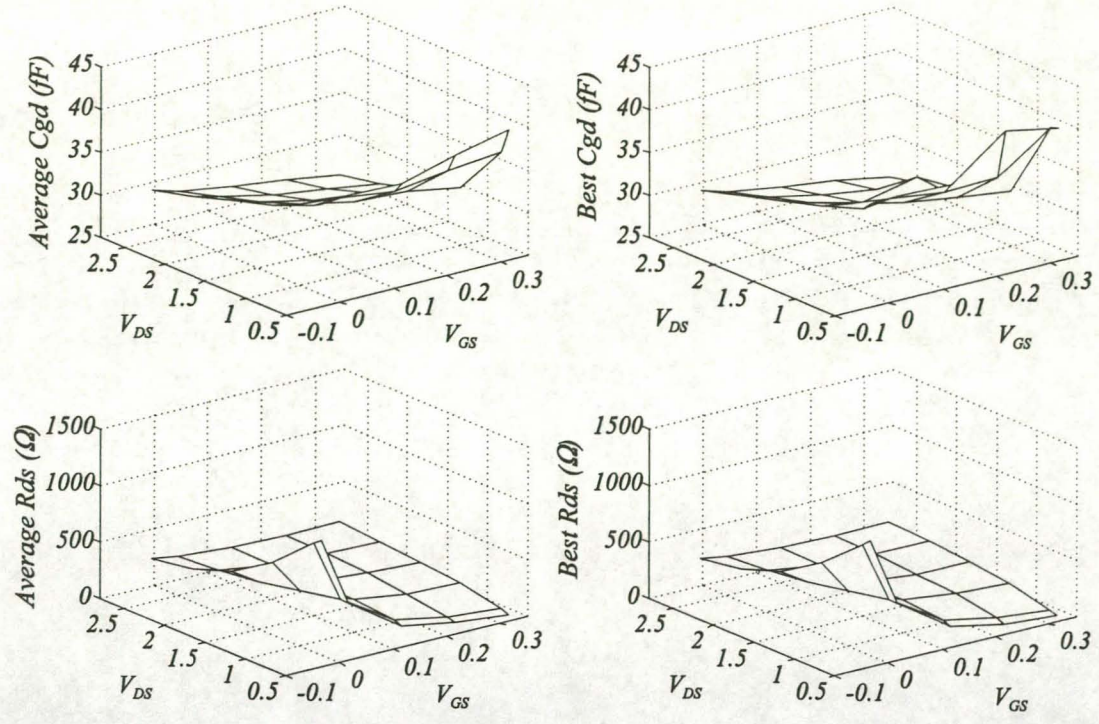


Fig. 2.36 The bias dependence of the pHEMT intrinsic elements Cgd and Rds.

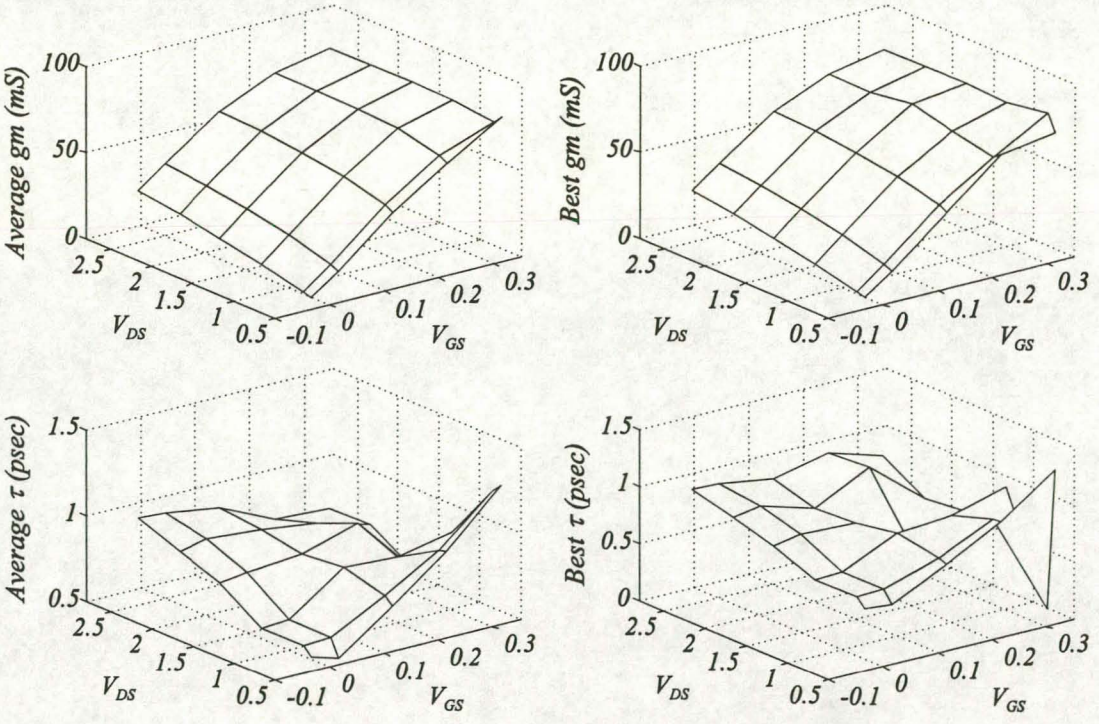


Fig. 2.37 The bias dependence of the intrinsic pHEMT elements gm and τ.

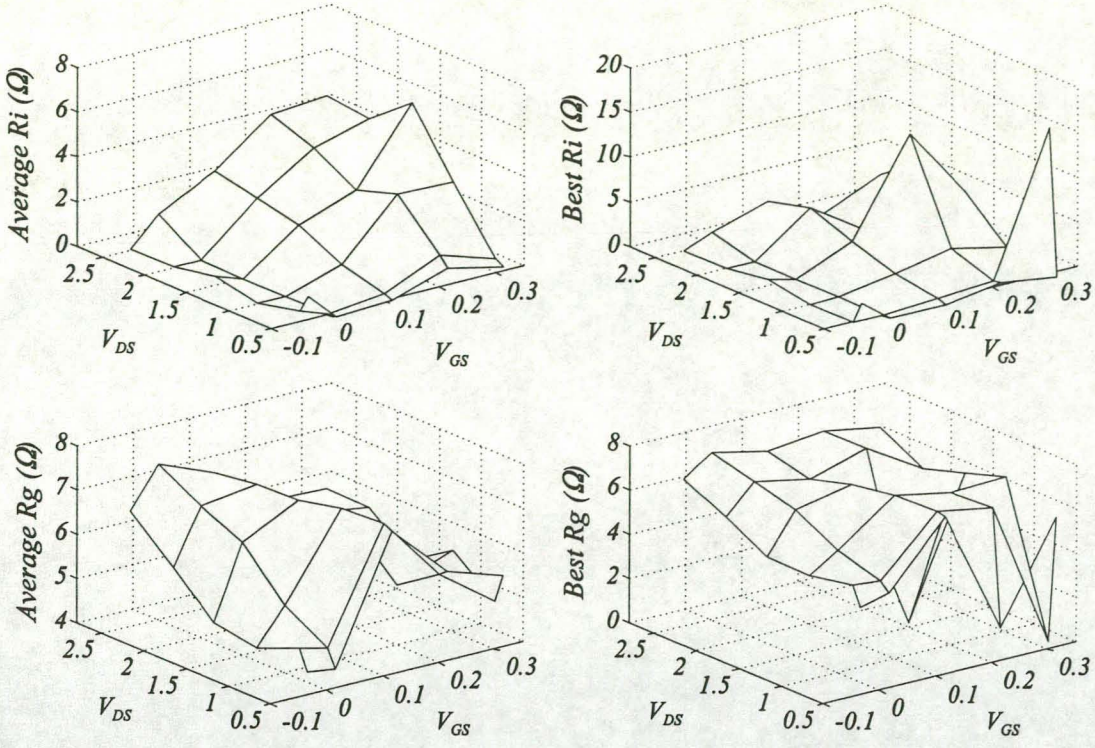


Fig. 2.38 The bias dependence of the pHEMT intrinsic element Ri and the extrinsic gate resistance Rg.

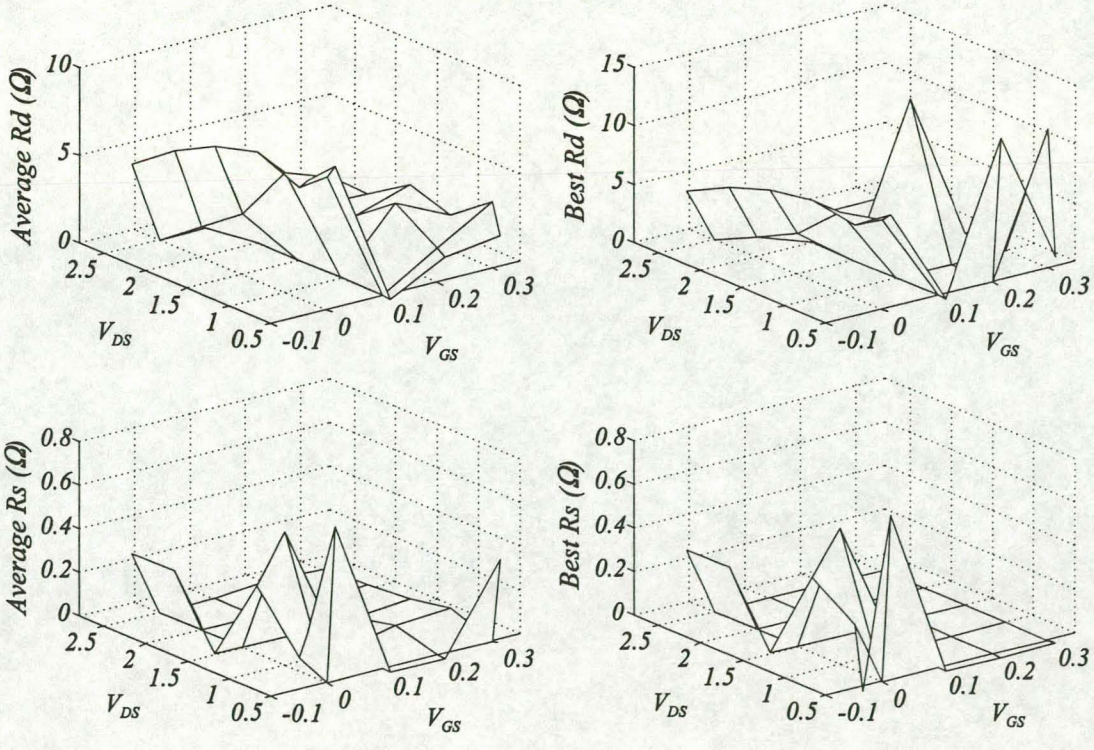


Fig. 2.39 The variation of the pHEMT extrinsic elements Rd and Rs with bias.

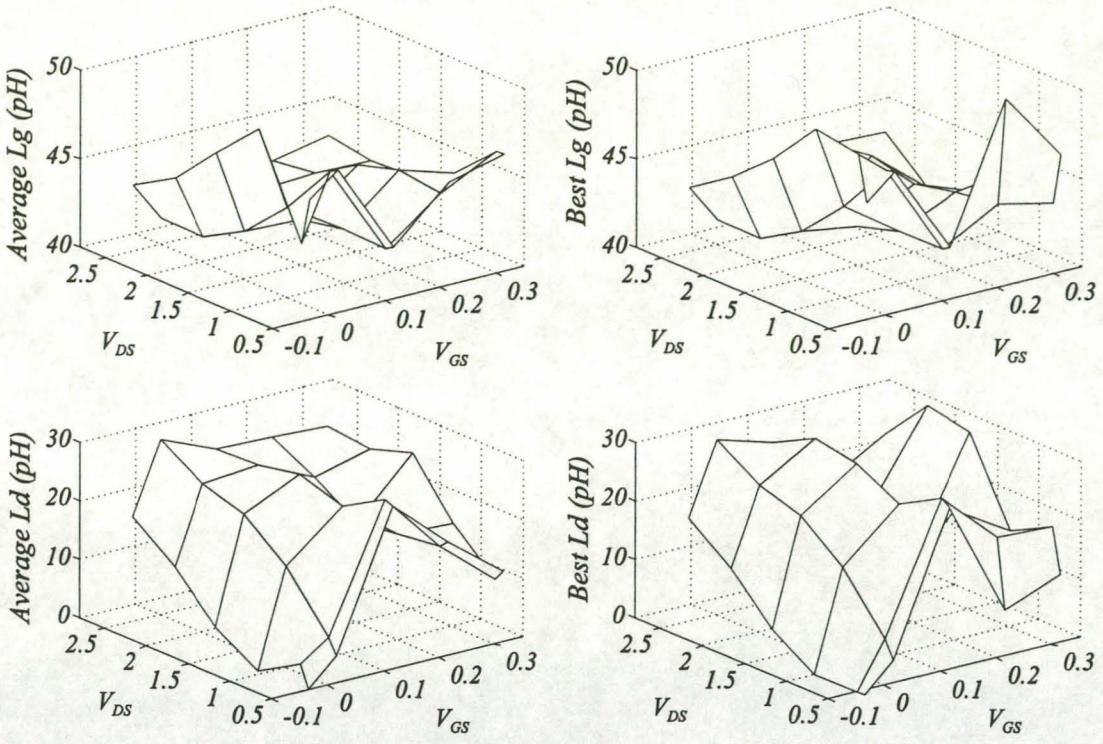


Fig. 2.40 The variation of the pHEMT extrinsic elements Lg and Ld with bias.

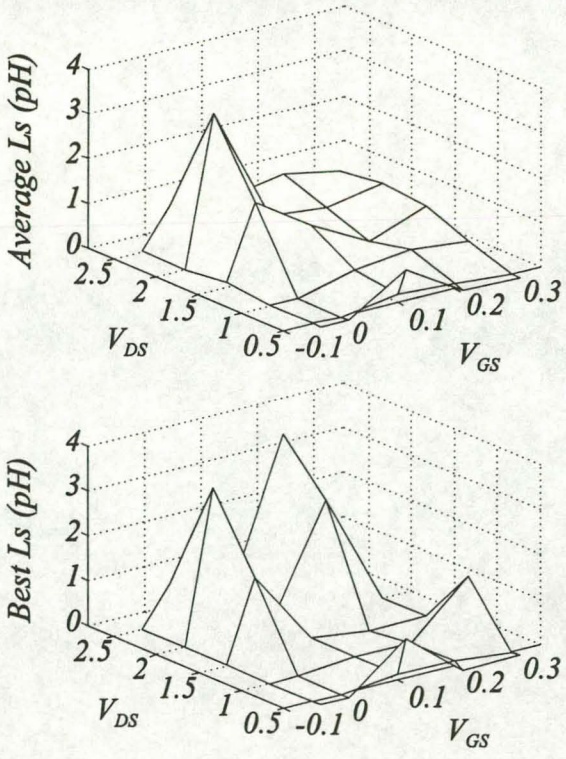


Fig. 2.41 The variation of the pHEMT extrinsic element Ls with bias.

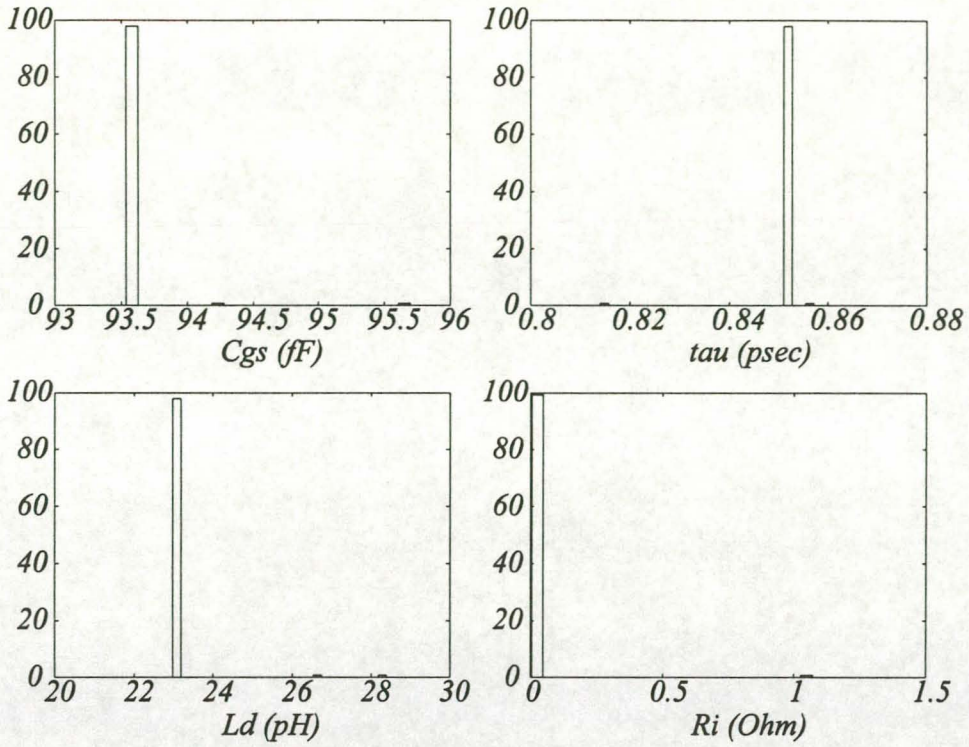


Fig. 2.42 The distribution of the pHEMT elements Cgs, τ, Ld and Ri after a robustness test at the bias point V_{GS} = 0.0 Volt and V_{DS} = 1.5 Volt.

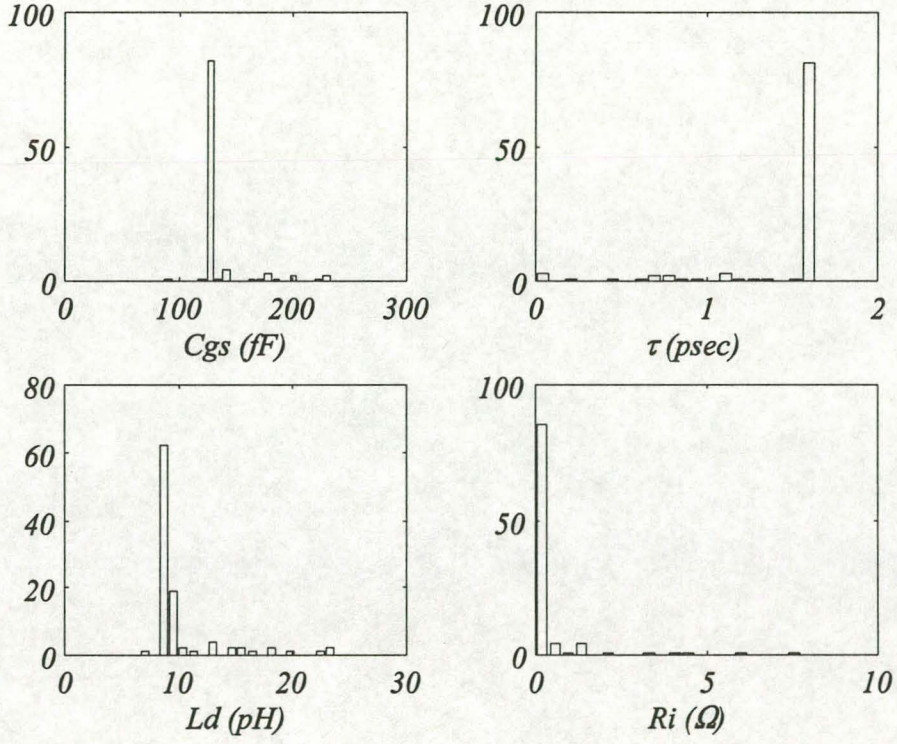


Fig. 2.43 The distribution of the pHEMT elements Cgs, τ, Ld and Ri after a robustness test at the bias point V_{GS} = 0.4 Volt and V_{DS} = 0.4 Volt.

The trends in the extraction results for the pHEMT are the same as for the MESFET, with the average and best element values being close to each other for the dominant elements at most of the bias points. The measured s-parameters of the pHEMT contain more imperfections (small resonances) than the MESFET data, and this is reflected in all the extracted element values. The variation of the elements with the applied bias voltages is not as smooth as for the MESFET. As with the MESFET, it is noted that in most cases the average element values are smoother functions of V_{GS} and V_{DS} than the best element values. The extrinsic elements show an even larger variation with bias than for the MESFET. The parasitic source resistance and inductance are particularly difficult to determine. Unlike the MESFET, the parasitic resistance Rd shows no clearly defined bias dependence.

Both the MESFET and the pHEMT exhibit small element distributions for bias points that are well within the saturated region of the device I_{DS}-V_{DS} bias plane. For the highly non-linear region near the knee-voltage, the distributions increase dramatically. It was also found that for bias points in or near the cut-off region of the transistors, the algorithm is more likely to terminate in a local minimum far from the optimum solution. For success at these bias points, the accuracy of the optimisation starting values become more important The results of single bias extractions are thus dependent on the choice of bias points used.

Figures 2.44 and 2.45 show the modelled and measured magnitude and phase of the pHEMT s-parameters at the bias point V_{GS} = 0.0 Volt and V_{DS} = 1.5 Volt. The average extracted elements were again used for the model-based s-parameter predictions. Like the MESFET, the differences between the s-parameters predicted with the average and best element values were negligible. The pHEMT shows the same trend as the MESFET in that the largest modelling errors occurred in the magnitude of the measured s-parameters. Extractions at other bias points exhibited similar modelling accuracies to those shown in figure 2.44 and 2.45.

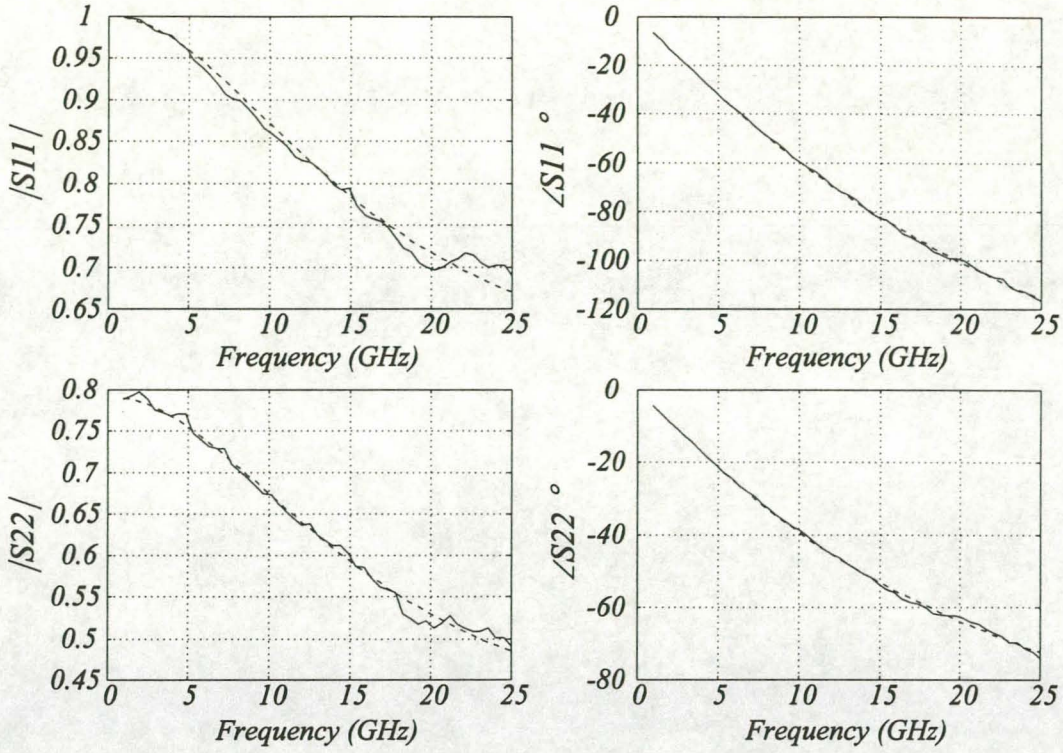


Fig. 2.44 A comparison of the modelled and measured pHEMT s-parameters at bias point V_{GS} = 0.0 Volt and V_{DS} = 1.5 Volt. (___ measurement, ---- model)

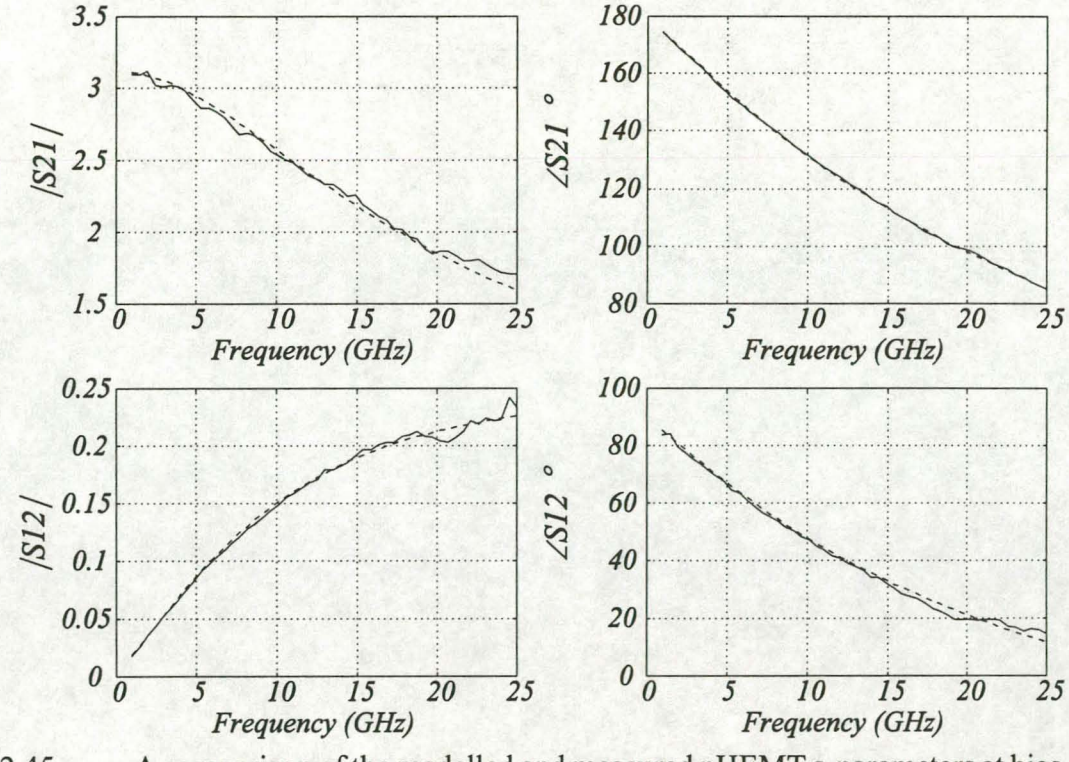


Fig. 2.45 A comparison of the modelled and measured pHEMT s-parameters at bias point V_{GS} = 0.0 Volt and V_{DS} = 1.5 Volt. (___ measurement, ---- model)

2.7 Convergence

No conclusive theoretical work is available to explain why the decomposed optimisation process converges to the correct solution. Kondoh [64] found that the order in which the functions are to be optimised is crucial for convergence. This is because the sub-optimisation problems are not linearly independent. In the work of Bandler and Zhang [77] it is merely stated that convergence is not guaranteed. In this section a closer look will be taken at the convergence behaviour of the extraction problem [79].

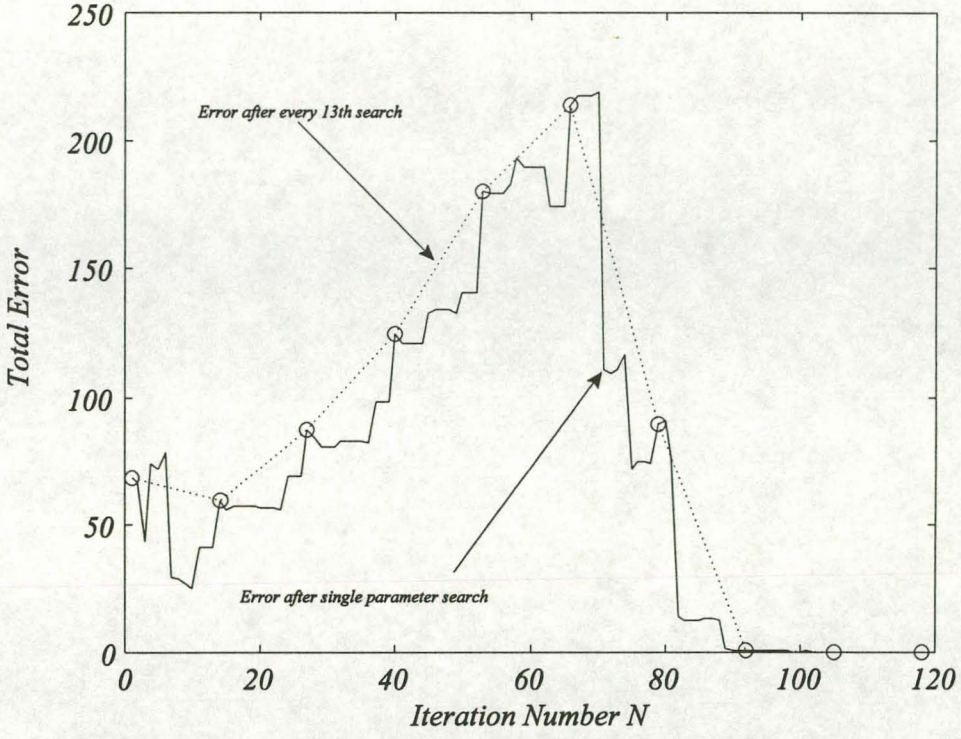


Fig. 2.46 The change in the global error (complete objective function) as the search progresses. The optimisation algorithm show signs of hill climbing. This example uses the FLR106XV transistor and s-parameters generated from the model.

The convergence of the algorithm can be divided into three regions. The third region of convergence is not visible in the simulated data, only when measured data is used. In the first region there occurs what can best be described as the pre-convergence manoeuvring of the model elements. During this phase the model element values change quickly, covering the whole search space in no apparent orderly fashion. This can be accompanied by large increases and decreases in the global error, giving the impression that the search routine is hill climbing. This behaviour is illustrated in figures 2.46 and 2.47.

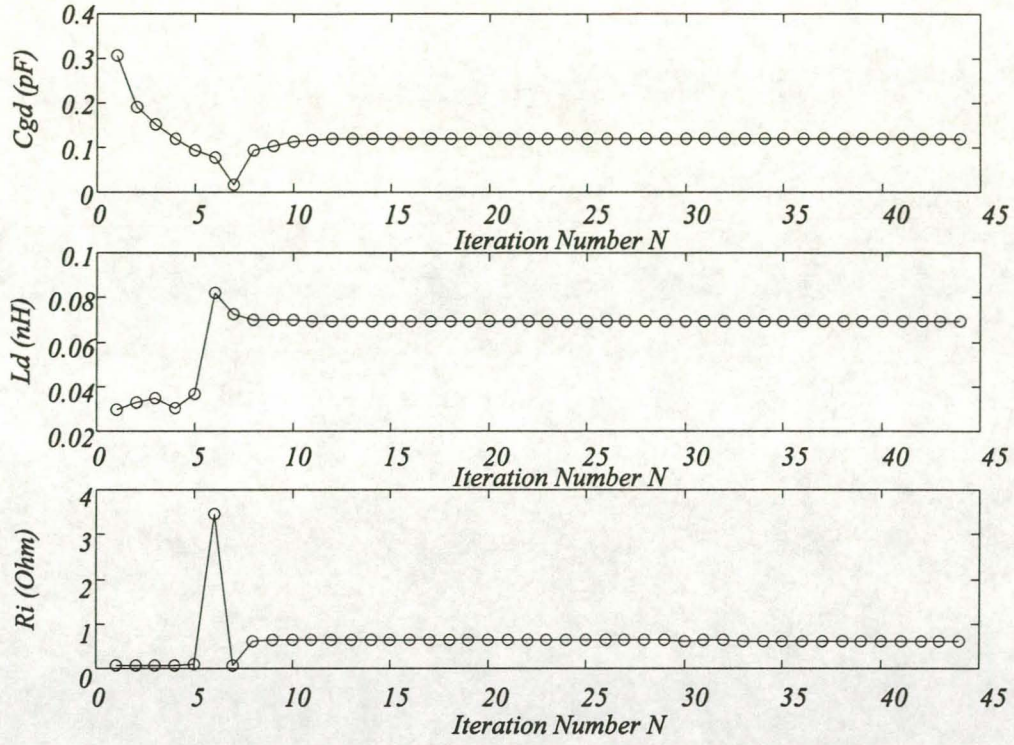


Fig. 2.47 The change in the element values of three elements of the FLR016XV 13 element model as the optimisation proceeds. The changes in the element values explain the apparent hill climbing of the search. Phase one and two of the algorithms convergence can also be seen.

One possible explanation is that the elements are shifted around until a point is reached from which the algorithm can converge to the final solution in an orderly fashion. Since the dominant model elements are initially far from their correct values, they will have a large influence on each other and the other less dominant model elements. This will increase the dependence the decomposed functions have on each other and explain the violent changes in the element and error function values.

The second phase of the algorithm's convergence is also evident from figures 2.46 and 2.47. During this part the algorithm converges smoothly to the region of the final solution. When the decomposed error functions are compared to the global objective function, it becomes evident that the subfunctions approximate the global objective function as the global minimum is approached. The convergence of the routine in phase two of the algorithm is therefore equivalent to the more traditional gradient-based optimisation routines. This whole process is illustrated graphically in figures 2.48 and 2.49.

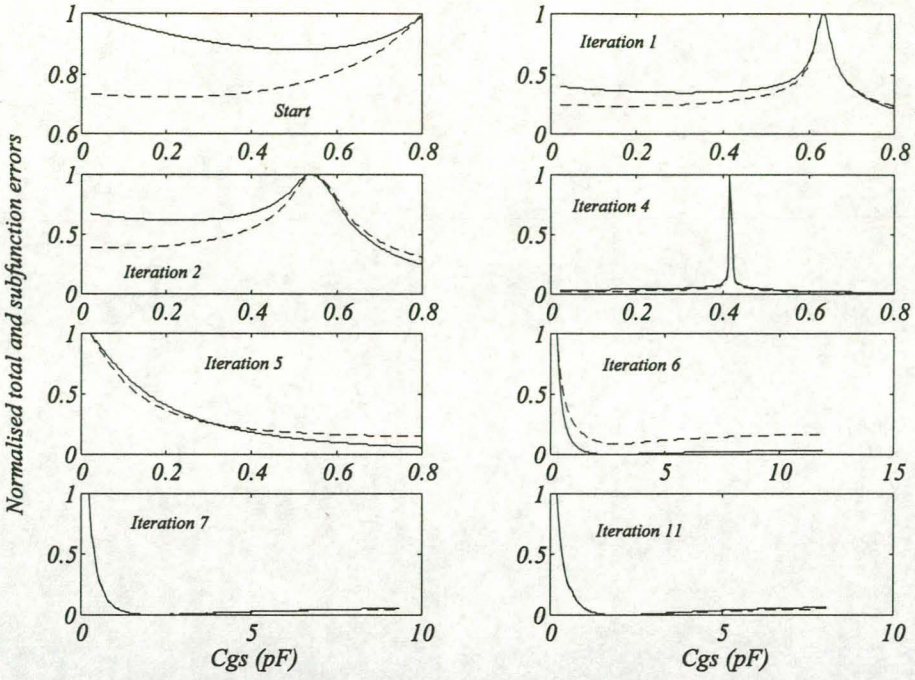


Fig. 2.48 The change in the landscape seen by optimiser with respect to the parameter Cgs. Snap shots of the value of Cgs is taken at different iterations of the search and the value of Cgs is swept around this value to provide the shape of the landscape that will be seen by the optimisation routine. The solid line is the subfunction error, while the broken line is the global error.

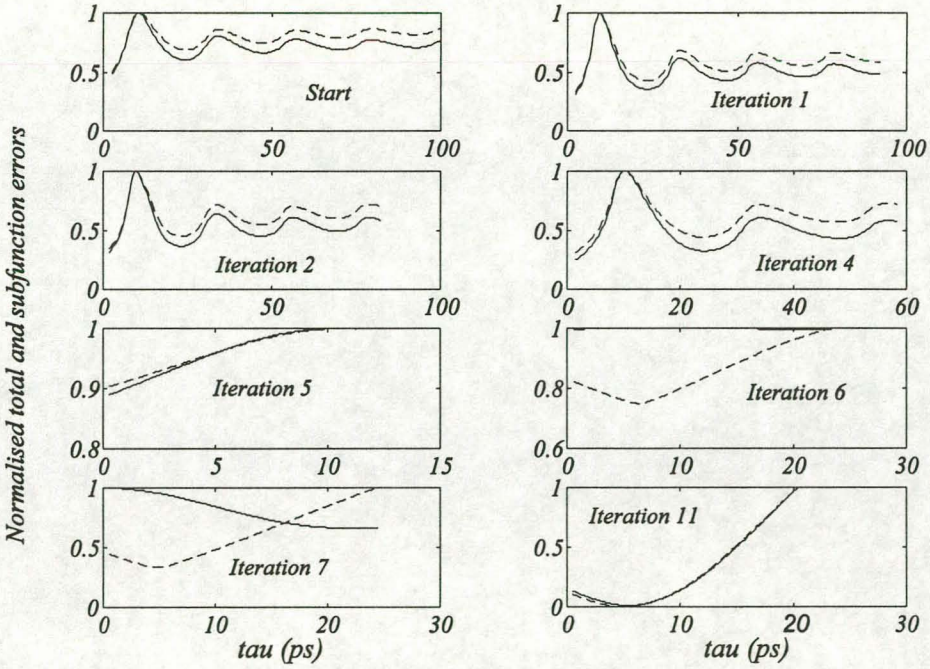


Fig. 2.49 The change in the landscape seen by optimiser with respect to the parameter τ. The solid line is the subfunction error, while the broken line is the global error.

Figures 2.48 and 2.49 clearly illustrate how the global objective function differs from the subfunction that is being optimised during phase one. The figures also show just how quickly the shape of the landscape can change from iteration to iteration, especially during the first few iterations (phase one of the search). Phase two of the algorithm's convergence is also clearly shown in the graphs. As the search continues, the global and subfunctions become better approximations of each other.

When dealing with measured data, a third phase of convergence is also evident. The algorithm converges to a minimum (end of second phase), after which it moves away from this solution and converges to a point close to the best solution that it found. An example of this is shown in figure 2.50 (a) for an extraction from MESFET data. Figure 2.50 (b) contains the convergence of the objective function when the extraction is performed using simulated data. The simulated measurements were generated using the MESFET model elements extracted in the experiment shown in figure 2.50 (a). The convergence towards a minimum is smooth when using simulated measurements.

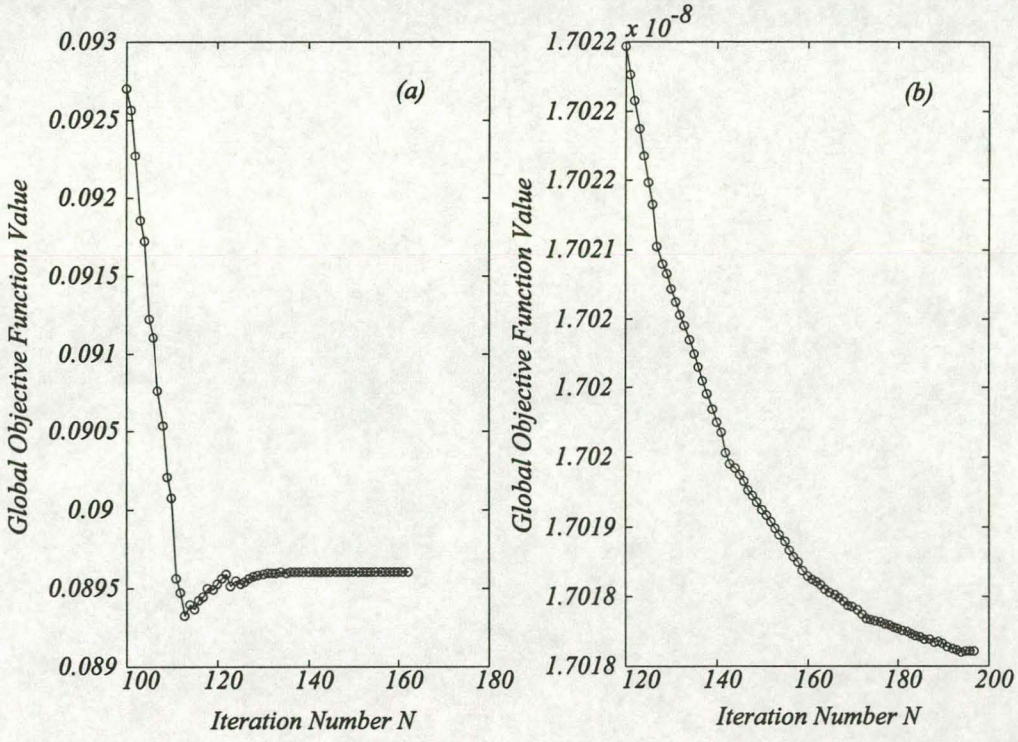


Fig. 2.50 A comparison of the convergence behaviour of the decomposition algorithm when (a) real measured data and (b) simulated s-parameters are used. The real data shows a third phase of the convergence where the objective function increases again before the algorithm converge to a stable solution. The measured data used in (a) are for the MESFET biased at V_{GS} = -0.6 and V_{DS} = 3.0 Volt.

The difference between the two experiments is that the simulated data allows for a perfect match between modelled and measured data. This implies that the solution of the global objective function and the subfunctions are the same. When using measured data, the model approximates reality, and there is no guarantee that the subfunctions and the global objective function will have exactly the same solution. Small differences between the optimum subfunction solutions will cause a tug of war between the subfunctions. The final solution to which the process converges will thus be a point at which the subfunctions are in balance with each other. This idea is further illustrated in figure 2.51 using simulated measurements. The simulated data is generated using the previously extracted MESFET elements and small parasitic capacitors that are added to the model. The data was further changed with the addition of small systematic random deviations that represent measurement biases [66]. These perturbations cause the optimum solution of the subfunctions and the global objective function to differ. The third phase of convergence now becomes visible.

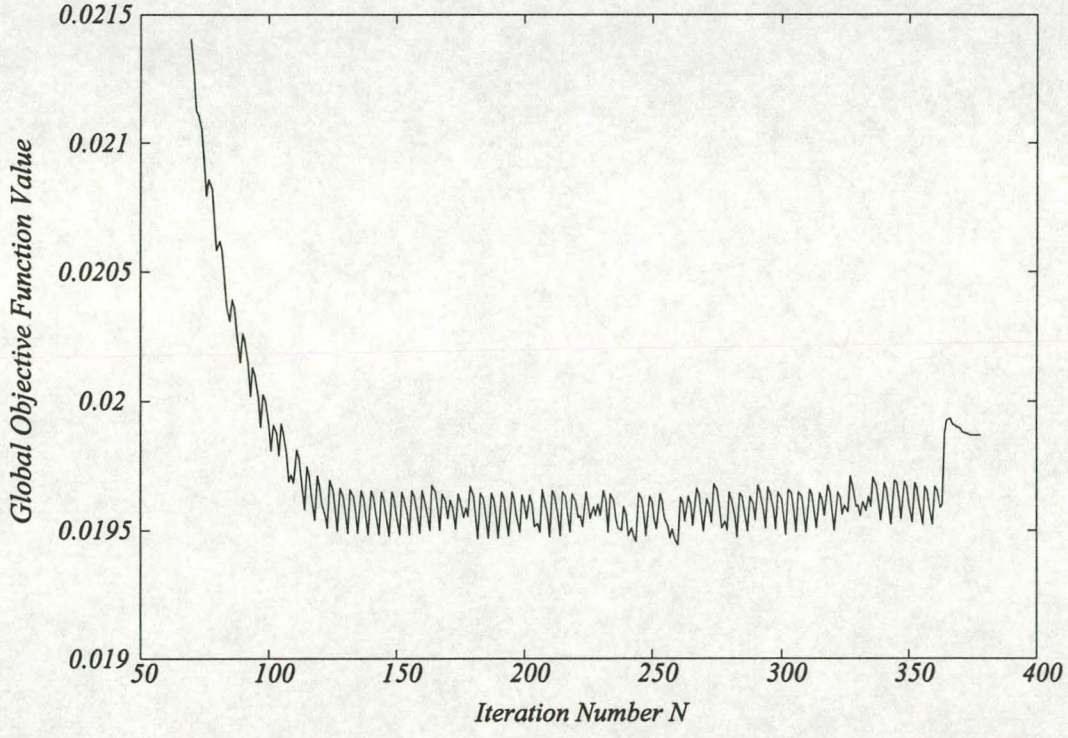


Fig. 2.51 The simulation of the third phase of the convergence behaviour using simulated measurements. The extracted MESFET elements were used as basis for the simulated data, with the effects of more parasitic elements and randomly chosen systematic deviations added to the simulated data.

The decomposed optimisation algorithms are by no means globally convergent and can still be caught in local minima if the search is started too far from the global minimum. The routines are however far more robust than most conventional search algorithms and are capable of covering a far larger search area. This robustness is believed to be due to several reasons:

- The error functions that are optimised in the search are far simpler than the global objective function, causing it to have fewer local minima. The global error is computed and stored during the optimisation process, but it is never used to make any decisions in the parameter extraction procedure.
- The measured data used in the extraction process contains quantities of greatly varying magnitude. Some measured s-parameters therefore have large differences in the standard deviation of errors made during their measurement. Conventional optimisation procedures that minimise the global error function cannot make any provision for this since the information is seldom available. This weakness is eliminated in the decomposition-based algorithm since the different measurement functions are never optimised together [65].

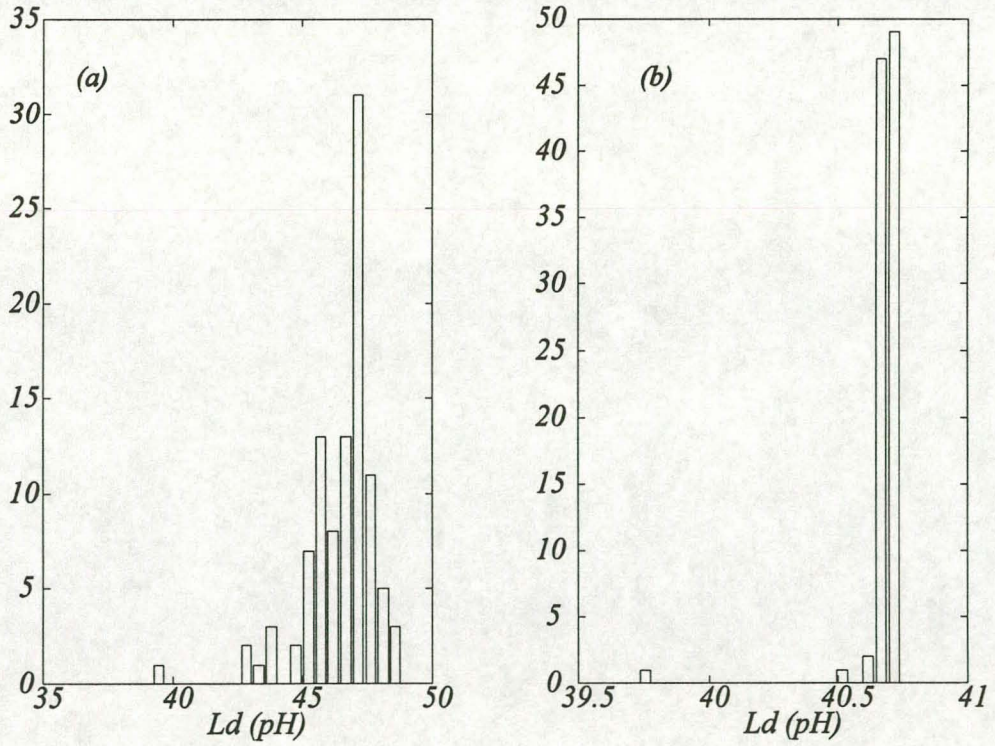


Fig. 2.52 A comparison of the spread in the extracted values of Ld when different termination conditions are used. The extractions were performed using the MESFET data. In (a) the global objective function was used as the stopping criteria, while in (b) the change in the values of the model elements was used.

The termination conditions for the decomposition-based search still needs to be stated. Given the complex behaviour exhibited by the optimiser, it is not possible to use the conventional approach that monitors the value and rate of improvement in the global objective function. A further complication is that the ill-conditioned nature of the problem makes it impossible to relate the changes in the objective function to changes in the values of the less dominant elements. Only the change in the values of the model elements is therefore monitored. Should all the model elements change by less than a certain percentage for a specified number of consecutive iterations, the algorithm is terminated. The model elements corresponding to the best solution found during the search are stored, and returned as the final extraction solution. These criteria cannot detect the third phase of the algorithm's convergence, and therefore allow for a certain number of wasted iterations. They are however very safe. This is considered far more important than efficiency.

A robustness test was performed in which the searches terminated if the global objective function changed by less than 0.01% for two or more consecutive iterations. Figure 2.52 compares the results obtained with that of the same search, but using the termination criteria described in the previous paragraph. The distribution of the drain inductance Ld is shown. It was found that the use of the global objective function to terminate the search resulted in the optimisation being terminated prematurely. This has a large influence on the uniqueness of the extracted solution, which is evident from the differences in the distribution of Ld shown in figure 2.52.

2.8 Pre- and Post-Processing Multi-Bias Parameter Extraction

The previous sections gave detailed experimental results describing the capabilities of the parameter extraction algorithm. It is evident that there is a limit to the experimental accuracy that can be achieved when the data at different bias points are considered individually. The next logical step would therefore be to find ways of combining the results of different single bias extractions.

The most commonly found method for doing this relies on the use of cold FET measurements, and can be referred to as a pre-processing approach. Optimisations using the cold FET data are performed to determine the extrinsic element values (pre-processing), which are then used as the starting values for extractions performed at other bias points. Since the decomposition-based extraction algorithm is not suited for use with cold FET models (section 2.5.4), this is not a viable multi-bias algorithm.

A post-processing approach can however be employed. This algorithm uses cold FET data to improve the element values extracted from hot s-parameter data. The decomposition optimiser can accurately determine values for most of the model elements from the hot s-parameter data. The extracted reactive extrinsic elements (Cpg, Cpd, Lg, Ld and Ls) are then de-embedded from the cold FET data, and analytical equations are used to calculate the values of the parasitic resistances (Rg, Rd and Rs) and the intrinsic capacitors. Once a value for Rg has been found, it is fixed, and the hot FET extraction is repeated. The method works on the assumption that if Rg is fixed close to its correct value, then the accuracy with which the other model elements can be determined will increase, which in return will allow a more accurate value for Rg to be estimated from the cold FET data.

In order to improve the chance of the post-processing algorithm converging, the initial hot extractions can be repeated a number of times using random starting values. Only the elements corresponding to the lowest modelling error are used. This makes the method more robust, but also inefficient. A further refinement that can be implemented is to perform extractions from several bias points and to make use of the average values of the extrinsic elements extracted. Good results were achieved with simulated measurements, but the addition of noise to the data caused the algorithm to fail. The success of the technique is dependent on its ability to determine an accurate estimate for Rg, an element whose value will be greatly influenced by measurement errors.

Both the pre- and post-processing approaches rely heavily on single bias extractions, and can therefore not be expected to do better than the single bias extraction, especially where the less sensitive elements are concerned.

2.9 Conclusions

This chapter describes the development of an adaptive decomposition-based optimisation algorithm for the extraction of FET small-signal models from measured s-parameters. This approach divides the problem into simpler subproblems which are solved iteratively in a specific sequence. Previous attempts used fixed sequences determined via experimentation [64], or made use of very heuristic algorithms [77]. The proposed algorithm uses a principal components sensitivity analysis to calculate the order of optimisation. An overview was given of the sensitivity analysis, its advantages and the factors that influence it.

The extraction algorithm was tested using both simulated and measured s-parameters. Tests

using simulated measurements provided information concerning the absolute accuracy of the procedure and its convergence behaviour. The results indicate that models with up to 16 elements can be extracted with reasonable accuracy. The only element that could not be determined was Rg. Novotny's [63] observations concerning the influence of Rg on the extraction accuracy were investigated and confirmed. Robustness tests using a wide range of random starting values were performed. These tests were repeated with the simulated measurements contaminated with noise. The results of the tests showed that the use of standard statistical descriptions for summarising the distributions of the extracted element values can be highly misleading. The factors that influence the distribution of the model elements were discussed.

The adaptive decomposition-based procedure was also compared with the decomposition algorithm proposed by Kondoh [64]. Using simulated measurements, it was determined that the new decomposition algorithm handles a very wide range of optimisation starting values better. Dominant, but traditionally difficult to determine elements such as τ, were not always extracted correctly by Kondoh's approach.

Measured data from two different devices, a GaAs MESFET and a GaAs pHEMT, were also used to evaluate the extraction procedure. The algorithm was subjected to similar robustness tests as was used with the simulated data. The measured s-parameters highlighted certain peculiarities in the convergence of the algorithm. The extraction results show that the algorithm converges to a tightly defined solution, independent of the optimisation starting values used. The uniqueness of the solution is however dependent on the bias point of the device. Near the knee region of both devices, a large increase in the distribution of the extracted elements is noted. This is a highly nonlinear region of the device's bias plane, with fast changes in the bias currents and voltages occurring. It was also found that at bias points near or in the cut-off region of the transistor, the routine is far more susceptible to being caught in a disastrous local minimum, depending on the optimisation starting values used.

In most cases it was found that the intrinsic elements are smooth functions of the applied bias voltages. The variation in the values of the extrinsic elements is however unacceptably high. The problem is that in order to achieve close fits between modelled and measured data, the model elements are following the imperfections in the measured s-parameters. These deviations in the data differ from bias point to bias point. This problem becomes even more clear when the extraction results of the pHEMT are compared with those of the MESFET. The pHEMT measurements contain far more small measurement errors (small resonances) than the MESFET data.

The next chapter describes a multi-bias decomposition-based optimiser that combines data from different bias points into a single extraction problem. This is done to more accurately determine the values of the insensitive model elements and to combat the influence of measurement imperfections.

Chapter 3
Multi-Bias Decomposition-Based Parameter Extraction

3.1 Introduction

As shown in the previous chapter, the single bias extraction cannot determine all the model elements, and determines the extrinsic elements with an unacceptably large variation with bias. To overcome this, several solutions that make use of data measured at different bias points, have been suggested [29,58,66,69,70,73]. These procedures exploit the fact that the extrinsic elements are bias independent.

The simplest multi-bias approach uses cold FET (V_{DS}=0) measurements with the gate voltage of the device held below its pinch-off voltage. For this cold bias, the intrinsic model of most FET devices reduces to only three capacitors. Since the model has fewer elements, it is argued that a simpler optimisation problem can be constructed to determine the values of the extrinsic model elements accurately. However, the cold FET is a passive circuit, thus also providing 25% less data from which to extract the model elements. This optimisation procedure is as ill-conditioned as the hot FET parameter extraction, as the large differences in objective function sensitivity with respect to the model elements, still prevails. The extrinsic elements are also determined from a limited set of data which includes measurement errors, and by fixing their values in subsequent optimisations, a propagation of errors will occur. Like the direct extraction methods, this procedure is an open loop, because there is no feedback that will attempt to correct errors made during any of the extraction steps.

In order to achieve a closed loop procedure, the s-parameters measured at different bias points have to be integrated into one combined extraction process. Current multi-bias extraction algorithms that do this, make use of conventional gradient optimisers [29,58,66,69,70] or random optimisers [73]. Bandler *et. al.* [58] described a multi-bias extraction that uses the l_1 norm. No assumptions were made concerning the bias dependencies of elements and any bias independent behaviour was reinforced with the aid of penalty functions. The technique relies heavily on the theoretical property of the l_1 norm to ignore large deviations. Lee [70] implemented a multi-bias extractor for HBT devices using a commercial frequency domain circuit simulator/optimiser. Two sets of s-parameters measured in the active mode of the device and one set of s-parameters measured with the device biased in the cut-off mode, were used. The extrinsic elements were taken to be bias independent and the current dependent elements in the active bias circuits were

linked to each other by the ratio of their currents. Cai *et. al.* [69] also implemented a multi-bias extraction technique for HBTs and made use of direct extraction techniques and a Levenberg-Marquardt optimiser. Patterson *et. al.* [29] proposed a multi-bias algorithm that relies on the physics-based equivalent circuit model proposed by Ladbrooke [27]. This model is better conditioned than the normal equivalent circuit small-signal model, but it is not valid for bias points outside the saturated region of V_{DS}-I_{DS} curves. It also needs process parameter values, which will be difficult to obtain for commercial devices. Patterson also made use of a principal components analysis and orthogonal transformations to further improve the conditioning of his optimisation procedure. Lin and Kompa [66] proposed a technique in which the extrinsic elements were optimised and the intrinsic elements were calculated using a new set of robust closed form equations. By extending this approach to multi-bias data, they showed that the uncertainty in the extraction of problematic elements such as Rg and Ri can be reduced greatly. Ghazinour [73] described a multi-bias extraction algorithm for MESFETs and HEMTs that makes use of a hybrid evolutionary/conjugate gradient optimisation approach. As random-based searches become very inefficient for large problems, Ghazinour reduced the number of variables in the problem by taking elements such as Ri, τ and Cds, which have a weak bias dependency in the active region, to be bias independent. The values of the remaining intrinsic elements were calculated as a function of the extrinsic elements.

All of these approaches optimise one global error function. This is inherently a difficult problem, as the number of local minima rises as the number of unknowns increase. To overcome problems related to the number of variables, current algorithms for solving the multi-bias extraction problem follow two distinct paths, namely:

- reducing the number of variables in the optimisation problem by representing the intrinsic elements as functions of the extrinsic elements with the aid of analytic expressions [66,73], or
- developing more robust optimisation methods that are suited to high dimensional problems [68,73,78].

In this chapter an integrated extraction method which uses s-parameter data measured at different operating bias points in the active region of a device, is presented. The technique falls within the second type of approach and is based on the adaptive decomposition-based optimiser discussed in chapter 2. The advantages of the proposed technique are:

- The decomposition-based optimiser does not make use of one global error function, making it more immune to local minima and the ill-conditioned nature of the problem

than other methods.

- The procedure combines the multi-bias data into one extraction process that determines all the element values simultaneously. This allows for tighter feedback than multi-bias methods based on single bias extractions, which in turn makes it easier for the algorithm to handle bad measurement points.
- Procedures that calculate certain elements with analytical equations [66,73], place a limit on the modelling fit that can be achieved, since these element values cannot be varied individually. The danger also exist of the analytical equations failing due to a few bad measurement points. The proposed procedure determines all the elements with optimisation and has proved itself to be resistant to the effect of measurement errors.
- Unlike some approaches, no assumptions are made concerning the bias dependencies of the intrinsic elements [73], and there is nothing in the algorithm that will suppress weak bias dependent behaviour in intrinsic elements [58].
- Since no assumptions are made concerning the device layout, the algorithm is suitable for modelling commercial devices about which very little additional information is known.

The multi-bias algorithm allows the extrinsic elements to be more uniquely defined, thus increasing the accuracy with which all the model elements are determined. The method has the same starting value independence as random searches, but it is far more efficient than a random based search. It only requires the user to provide optimisation boundaries, and as will be shown, the allowed optimisation search space can be made very large.

3.2 The Multi-Bias Objective Function

A new objective function has to be formulated that combines the measured s-parameter data and the model responses for the different bias points into one equation. The objective function is again formulated as the sum of the squared errors. This allows the decomposition and principal components analysis discussed in chapter 2, to be directly applied to the multi-bias problem.

The multi-bias objective function is defined as,

$$F(\bar{x}) = \sum_{i=1}^{N} \alpha_i(\bar{x})^2 \tag{3.1}$$

where

$$\alpha_i(\bar{x}) = \sum_{t=1}^{M}\sum_{j=1}^{2}\sum_{k=1}^{2}\frac{1}{\sigma_{jk}(t)}\left|R_{jk}(\bar{x}_{ind},\bar{x}_t,\omega_i)-S_{jk}(t,\omega_i)\right| \tag{3.2}$$

and

$$\sigma_{jk}(t) = \left|S_{jk}(t,\omega_i)\right|_{MAX} \tag{3.3}$$

In eq. (3.2) the symbols are defined as,

$\alpha_i(\bar{x})$	is the sum of the modelling errors made at frequency ω_i for all the s-parameters at all the bias points that are considered in the extraction,
N	is the number of frequencies at which measurements were made,
M	is the number of bias points in the extraction,
t	indicates the specific bias point that is being evaluated,
j, k	are the indexes for the four measured s-parameters,
$\bar{x}_{ind}$	is a vector containing all the bias independent model elements and
$\bar{x}_t$	is a vector containing the bias dependent elements describing bias point number t.

The rest of the symbols are the same as for the single bias extraction. A vector $\bar{x}$ can be defined that contain all the model elements for the multi-bias extraction and has the form

$$\bar{x} = [\bar{x}_{ind}\ \bar{x}_1\ \cdots\ \bar{x}_{M-1}\ \bar{x}_M] \tag{3.4}$$

This formulation of the model element vector is compact and it makes the evaluation of the Jacobian matrix fairly easy.

3.3 The Formulation of the Decomposition-based Multi-Bias Optimiser

The application of the decomposition process to the multi-bias optimisation problem is straightforward and follows the same steps that were outlined in chapter 2.

3.3.1 Subfunctions and Model Element Assignments

The multi-bias global objective function must be broken up into suitable subfunctions for use

in the decomposition-based optimiser. Since the extrinsic elements of the model are bias independent, their effect will be common to all the data. The multi-bias measurements therefore contain redundant information that can be used to define the extraction problem more uniquely. New subfunctions are created for the bias independent elements that make use of the information redundancy. By increasing the amount of extra information in the extraction problem, the determination of the extrinsic elements becomes more accurate. In the single bias extraction, certain elements are highly correlated because they look electrically the same. There is not enough information to distinguish between the elements according to their function. The multi-bias approach allows for the differentiation between extrinsic and intrinsic elements according to the function they perform. This reduces the inter element correlations and increases the accuracy with which the intrinsic elements can be determined.

In the decomposition-based single bias extraction, each model element is associated with the error made in modelling a specific s-parameter, and an error subfunction is defined for each element as

$$f_{jk}(x_l,\bar{x}_R) = \sum_{i=1}^{N} |R_{jk}(x_l,\bar{x}_R,\omega_i)-S_{jk}(\omega_i)|^2 \qquad (3.5)$$

where x_l is the model element being optimised and $\bar{x}_R$ contain the other model elements.

To include all the bias points in the extraction of the bias independent elements, the definition of the error subfunction for a bias independent element is changed to eq. (3.6).

$$f_{jk}(x_l,\bar{x}_R) = \sum_{i=1}^{N}\left(\sum_{t=1}^{M}\frac{1}{\sigma_{jk}(t)}|R_{jk}(x_l,\bar{x}''_{ind},\bar{x}_t,\omega_i)-S_{jk}(t,\omega_i)|\right)^2 \qquad (3.6)$$

where $\bar{x}''_{ind}$ contains the bias independent elements, with the exclusion of element x_l.

Eq. (3.6) represents the modelling error made in s-parameter S_{jk} over all the bias points and frequencies considered in the extraction. The use of eq. (3.6) for optimising the bias independent elements, creates four additional subfunctions that can be used in the decomposition process. Table 3.1 provides a summary of the multi-bias error function/model element associations that were used. The validity of these associations has been confirmed using the sensitivity analysis that was discussed in section 2.3.

Table 3.1 The decomposition of the 16 element FET model for the multi-bias extraction problem. Provision is made for the intrinsic feedback resistor Rj and the parasitic capacitors Cpg and Cpd.

	FET Model Parameter	S-Parameter Subfunction
1	Cgs, Ri	$f_{11}(x_l,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{11}(x_l,\bar{x}_R,\omega_i) - S_{11}(\omega_i)\rvert^2$
2	Cgd, Rj	$f_{12}(x_l,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{12}(x_l,\bar{x}_R,\omega_i) - S_{12}(\omega_i)\rvert^2$
3	Cds, Rds	$f_{22}(x_l,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{22}(x_l,\bar{x}_R,\omega_i) - S_{22}(\omega_i)\rvert^2$
4	gm, τ	$f_{21}(x_l,\bar{x}_R) = \sum_{i=1}^{N} \lvert R_{21}(x_l,\bar{x}_R,\omega_i) - S_{21}(\omega_i)\rvert^2$
5	Cpg, Lg, Rg	$f_{11}(x_l,\bar{x}_R) = \sum_{i=1}^{N}\left(\sum_{t=1}^{M} \frac{1}{\sigma_{11}(t)} \lvert R_{11}(x_l,\bar{x}''_{ind},\bar{x}_t,\omega_i) - S_{11}(t,\omega_i)\rvert\right)^2$
6	Cpd, Ld, Rd	$f_{22}(x_l,\bar{x}_R) = \sum_{i=1}^{N}\left(\sum_{t=1}^{M} \frac{1}{\sigma_{22}(t)} \lvert R_{22}(x_l,\bar{x}''_{ind},\bar{x}_t,\omega_i) - S_{22}(t,\omega_i)\rvert\right)^2$
7	Ls, Rs	$f_{12}(x_l,\bar{x}_R) = \sum_{i=1}^{N}\left(\sum_{t=1}^{M} \frac{1}{\sigma_{12}(t)} \lvert R_{12}(x_l,\bar{x}''_{ind},\bar{x}_t,\omega_i) - S_{12}(t,\omega_i)\rvert\right)^2$

3.3.2 Principal Components Sensitivity Analysis

The only difference in the principal components analysis when applied to the multi-bias objective function is in the way the Jacobian matrix is computed. The Jacobian matrix is defined similarly to eq. (2.13).

$$J(\bar{x}) = \begin{bmatrix} \frac{\partial \alpha_1(\bar{x})}{\partial x_1} & \frac{\partial \alpha_1(\bar{x})}{\partial x_2} & \cdots & \frac{\partial \alpha_1(\bar{x})}{\partial x_j} \\ \frac{\partial \alpha_2(\bar{x})}{\partial x_1} & \frac{\partial \alpha_2(\bar{x})}{\partial x_2} & \cdots & \frac{\partial \alpha_2(\bar{x})}{\partial x_j} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \alpha_N(\bar{x})}{\partial x_1} & \frac{\partial \alpha_N(\bar{x})}{\partial x_2} & \cdots & \frac{\partial \alpha_N(\bar{x})}{\partial x_j} \end{bmatrix} \tag{3.7}$$

By rearranging the order of the summations in eq. (3.2) and substituting eq. (2.14) into eq. (3.2), $\alpha_i(\bar{x})$ can be written as,

$$\alpha_i(\bar{x}) = \sum_{t=1}^{M} e_i(\bar{x}_{ind}, \bar{x}_t) \tag{3.8}$$

If x_l is a bias dependent element describing bias point t, then the partial derivative of eq. (3.8) can be written as shown in eq. (3.9)

$$\frac{\partial \alpha_i(\bar{x})}{\partial x_l} = \sum_{t=1}^{M} \frac{\partial e_i(\bar{x}_{ind}, \bar{x}_t)}{\partial x_l} = \left. \frac{\partial e_i(\bar{x}_{ind}, \bar{x}_t)}{\partial x_l} \right|_{t=1,2,\ldots,M} \tag{3.9}$$

The summation in eq. (3.8) falls away since element x_l only affects one bias point and the other $e_i(\bar{x}_{ind}, \bar{x}_t)$ terms are therefore seen as constants by the partial derivative.

If x_l is bias independent, it affects the modelling errors at all the bias points and the partial derivative of eq. (3.8) becomes as shown in eq. (3.10).

$$\frac{\partial \alpha_i(\bar{x})}{\partial x_l} = \sum_{t=1}^{M} \frac{\partial e_i(\bar{x}_{ind}, \bar{x}_t)}{\partial x_l} \tag{3.10}$$

Once the Jacobian matrix is known, the Hessian matrix can be approximated and the principal components analysis follows as before. From a programming point of view, the calculations used in the multi-bias extraction are no more complex than for the single bias problem. The only extra complication is that more bookkeeping needs to be done in order to distinguish between the bias dependent and bias independent elements.

Because of the large number of elements in the multi-bias extraction, it is not practical to list calculated optimisation sequences in a table. Figure 3.1 provides a graphical alternative. The y-axis represents the 15 elements of the model in figure 2.19 (a), while the x-axis shows their position in the optimisation sequence. Figure 3.1 indicates that Cpg is optimised first, followed by Cpd and the different values of Cgd etc. It is not possible to see from figure 3.1 to which bias points the different model elements belong, but the graph clearly shows the patterns in the optimisation order that was discussed in section 2.4.3. Also note that the dominant model elements for the different bias points group together. The previously insensitive element Rg has moved higher up the optimisation order. The less dominant model elements (such as Rds) are not always grouped together. This is to be expected since one of the factors influencing the objective function sensitivity to these elements are their values, which is dependent on the bias.

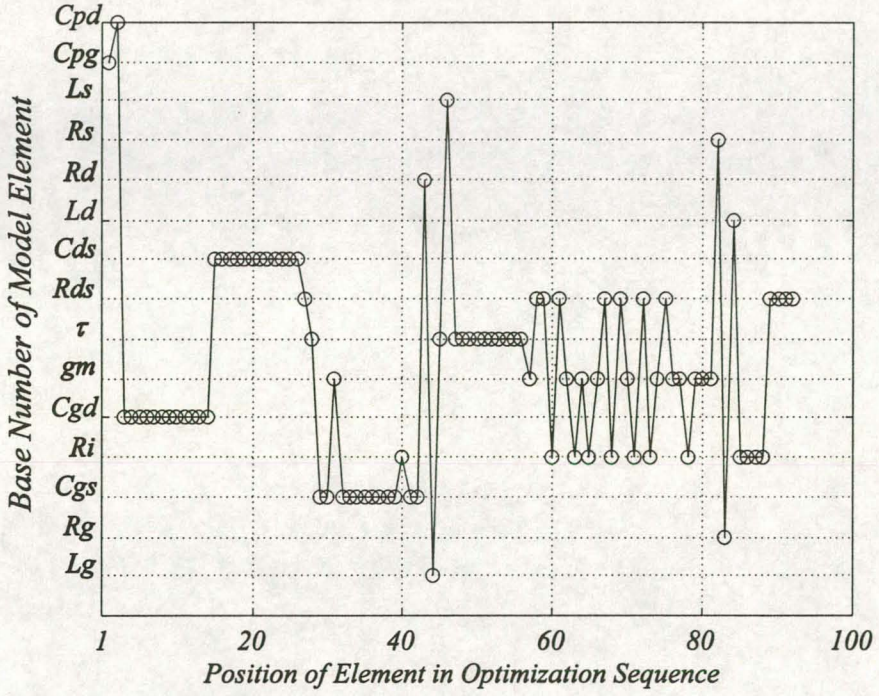


Fig. 3.1 A graphical illustration of the optimisation sequence generated by the multi-bias principal components analysis. The data was generated using simulated measurements and the 15 element model of figure 2.19 (a).

3.4 Experimental Results: GaAs MESFET and pHEMT Modelling

The multi-bias modelling algorithms were tested on the measured data of the MESFET and pHEMT devices used in chapter 2, using both 13 and 15 element models. Chapter 2 showed that the 15 element model leads to physically unrealistic model elements and non-unique extraction solutions. When compared to the single bias results, the 15 element model extracted from the MESFET data, using a multi-bias approach, results in more realistic model element values. However, the 15 element model does not provide a lower modelling error than the 13 element model. The 15 element pHEMT model, like the MESFET, failed to provide unique solutions. Modelling errors for the 13 element and 15 element pHEMT models were approximately the same, indicating that the parasitic capacitors have a larger influence on the pHEMT device. Direct measurements indicate that the values of the parasitic capacitors are between 10 and 15 fF, which is not small when compared with the values of the intrinsic capacitors Cgs and Cds [18,95,96]. The frequency range of the measurements is inadequate for uniquely separating the values of these elements.

In order to determine the robustness of the multi-bias procedure, a search using five bias points was repeated 100 times using randomly chosen starting values. Table 3.2 summarises the results for the MESFET and the pHEMT. The table shows the mean extracted value of each element. An uncertainty factor Δ is defined as the difference between the maximum and minimum extracted element values. Due to space constraints only the intrinsic elements for the bias point that had the largest modelling error are listed in table 3.2.

It is clear from table 3.2 that most of the model elements converge consistently to a very small range of values. This is also true for the traditionally insensitive elements such as Rg, Ri and τ. For the MESFET, the method could not find a value for Rd. The single bias extractions indicated that Rd has a very large bias dependence (figure 2.27), with a large number of bias points at which Rd wants to acquire a negative value. This large spread of Rd values, with most of them negative, explains why the Rd value extracted by the multi-bias routine always stayed at the minimum optimisation boundary. Rd and Rs are partly formed by the contact resistance of the metal semiconductor interface, and the resistance of the bulk GaAs between the contacts and the channel. For high drain voltages the depletion region under the gate extends considerably towards the drain contact, and the region of bulk GaAs contributing to Rd varies. The high fields that exist in the gate-drain region may also cause the electrons in that region to gain enough energy to make the transition to a higher satellite energy valley. This will cause the electrons to become heavier thus decreasing their velocity. The effect of this is to make Rd appear negative because of the formation of a localised Gunn domain. Rs is expected to be

constant since the shape of the depletion region varies little at the source side of the gate and there are no large fields in this region. While this effect has been confirmed with Monte Carlo simulations [28], much debate concerning the validity of negative parasitic resistances still remains [97].

Table 3.2 The extraction results for the multi-bias robustness test for the two different devices considered. (Δ = Max. Extracted Value - Min. Extracted Value)

	MESFET				GaAs pHEMT			
	Search Space		Results		Search Space		Results	
	Min. Value	Max. Value	Mean Value	Δ	Min. Value	Max. Value	Mean Value	Δ
Cgs (fF)	3.5E-2	1750	558.311	8.248	1.9E-2	950	66.132	0.0481
Cgd (fF)	4.4E-2	220	15.55	0.261	3.09E-3	154.5	36.698	0.0158
Cds (fF)	9.0E-3	450	71.306	0.327	1.28E-2	640.5	29.789	0.0628
gm (mS)	5.4E-3	270	66.185	1.074	5.0E-3	250	4.095	0.0027
τ (psec)	6.0E-4	30	3.852	0.0013	1.0E-4	5	0.864	0.0156
Ri (Ω)	1.5E-4	7.5	2.262	0.746	5.0E-4	25	1.37	0.246
Rds (kΩ)	2.56E-5	1.28	0.496	0.0081	1.0E-4	5	2.223	0.0317
Lg (pH)	9.0E-3	450	31.171	0.048	5.1E-3	255	43.546	0.156
Ld (pH)	9.0E-3	450	40.679	1.007	4.0E-3	200	17.326	0.677
Ls (pH)	2.0E-3	100	4.785	0.112	1.5E-4	7.5	0.1882	0.322 (2)
Rg (Ω) (1)	1.5E-4	7.5	1.702	0.59	1.5E-4	7.5	5.743	0.123
Rd (Ω)	4.4E-4	22	4.4E-4	0 (2)	3.0E-4	15	4.483	0.252
Rs (Ω)	2.9E-4	14.5	1.179	0.239	2.0E-4	10	2.0E-4	0 (2)

(1) If the worst one of the random extractions for the MESFET is discarded, then the Δ value for Rg changes to 0.1605 Ω, a reduction by a factor of 3.68

(2) See text

The pHEMT results show that all the model elements, with the exception of the source inductance and the source resistance, are extracted with a high degree of certainty. The source inductance varies considerably from extraction to extraction. The pHEMT measurements contain far more imperfections than the MESFET data, which could account for this, especially since Ls has such a small value. Rs is not a dominant element, and could also be affected by

measurement imperfections. Rs behaved similarly when it was extracted using the single bias algorithm (see figure 2.39) and it is therefore unlikely that the behaviour of Rs is due to bias dependencies.

The multi-bias extraction algorithm increases the uniqueness of the bias independent elements by increasing the amount of redundant information available to the optimiser for determining the extrinsic elements. The question of how many bias points are required for the model elements to be uniquely defined, needs to be answered. Robustness tests, each using an increasing number of bias points, were performed to investigate this. Figures 3.2 to 3.5 show the curves for extrinsic model elements and the intrinsic elements Cgs, Cds, Ri and τ as a function of the number of bias points used in the extraction. Extractions with up to 48 bias points were performed. Because of the large number of bias points, it was not practical to do robustness tests for all the multi-bias extractions. Robustness tests were performed for the 1,2, 5 and 10 bias point extractions. The curves shown in the figures represent the element values extracted from one random starting value. In figures 3.6 to 3.9 the Δ values for the different elements are plotted as a function of the number of bias points. The uncertainty factor is expressed as a percentage of the mean element values.

The results indicate relatively large variations in the element values if less than 5 bias points are used in the extraction. The extraction uncertainty also decreases rapidly for extractions using fewer than 5 bias points, and then stabilises. It is therefore concluded that 5 is the minimum number of bias points that must be used in the multi-bias algorithm if results that are independent of measurement errors are to be obtained. It is noted that the uncertainty Δ of the non-dominant element Ri is still high - approximately 50% for the MESFET and between 20% and 30% for the pHEMT. The accuracy with which these elements can be extracted will always be limited by the measured data due to their sensitivity to measurement imperfections [98].

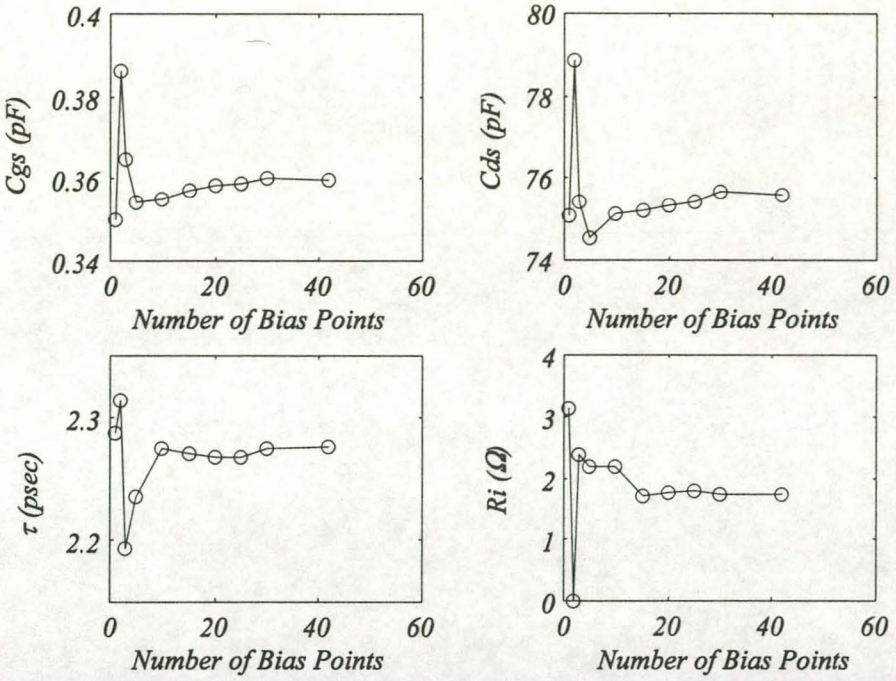


Fig. 3.2 The variation of the MESFET intrinsic elements as a function of the number of bias points used in the multi-bias extraction. These elements describe the bias point V_{GS} = -0.6 Volt and V_{DS} = 3.0 Volt.

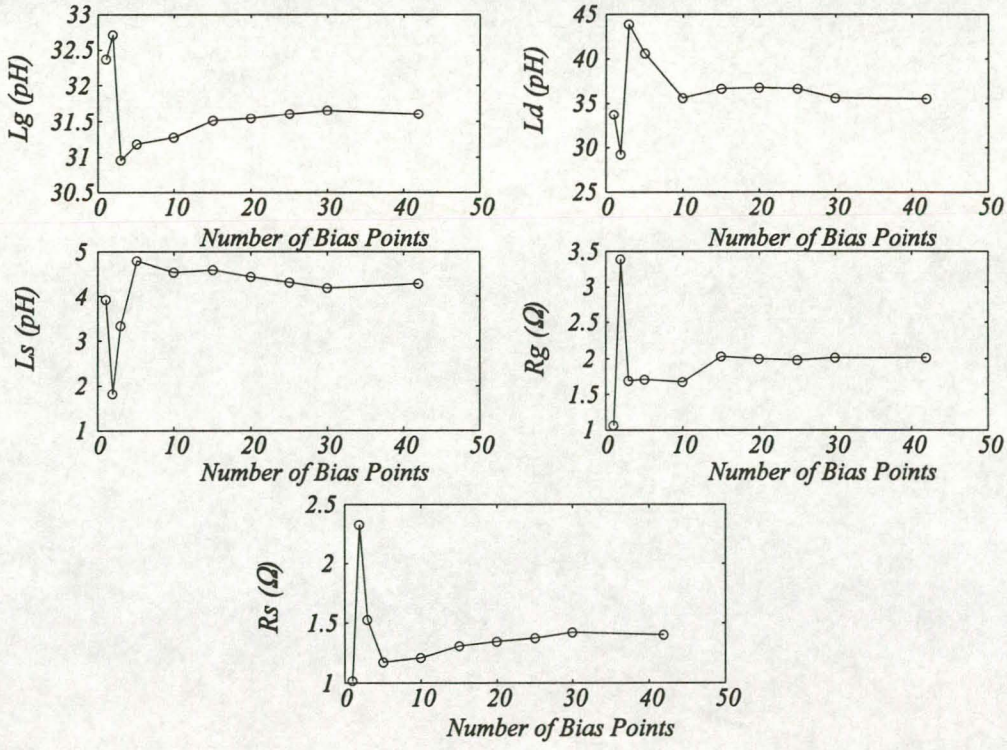


Fig. 3.3 The variation of the MESFET extrinsic elements as a function of the number of bias points used in the multi-bias extraction.

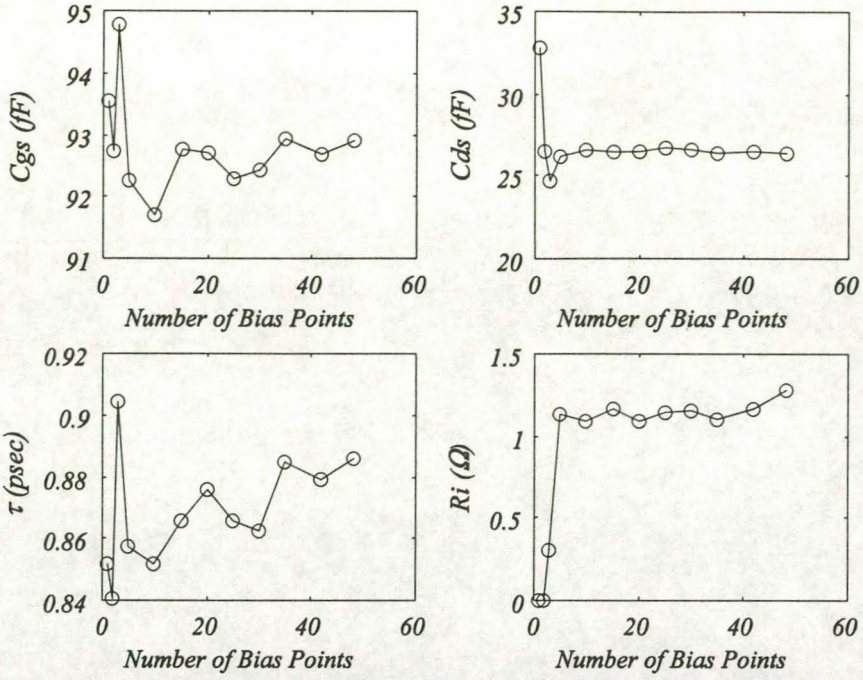


Fig. 3.4 The variation of the pHEMT intrinsic elements as a function of the number of bias points used in the multi-bias extraction. These elements describe the bias point V_{GS} = 0.0 Volt and V_{DS} = 1.5 Volt.

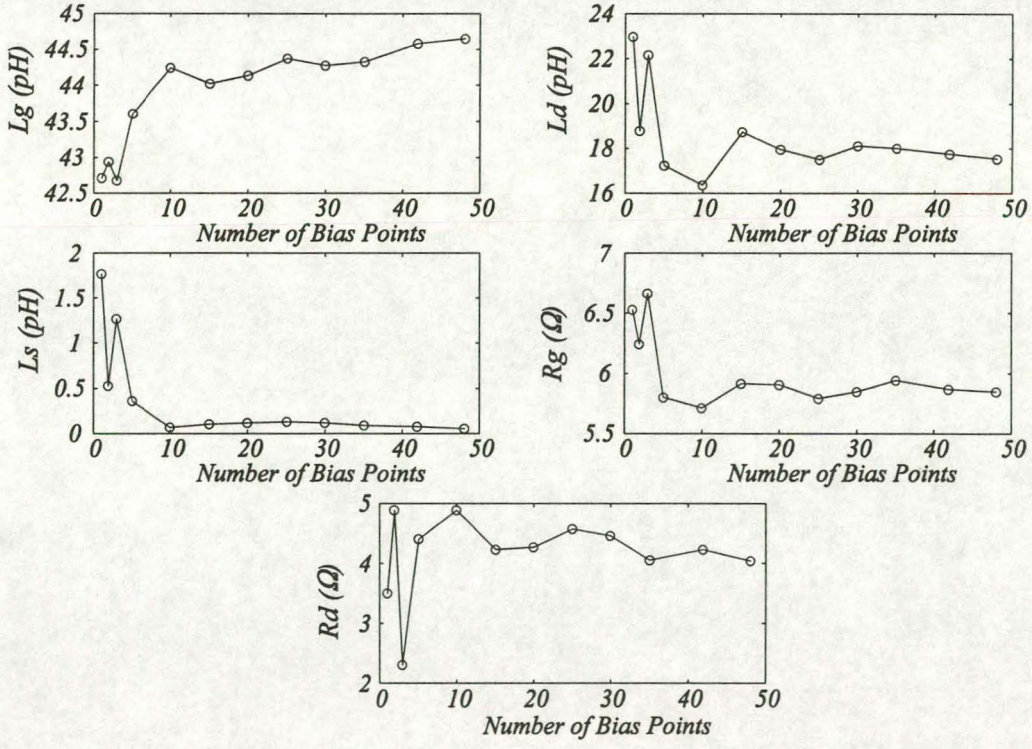


Fig. 3.5 The variation of the pHEMT extrinsic elements as a function of the number of bias points used in the multi-bias extraction.

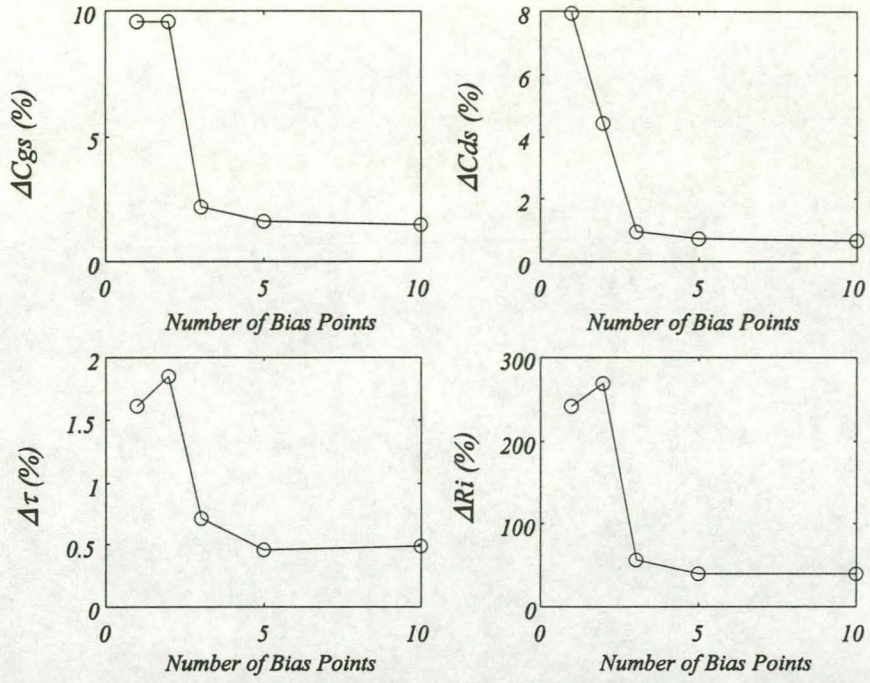


Fig. 3.6 The uncertainty in the MESFET intrinsic elements as a function of the number of bias points used in the extraction. The intrinsic elements are for the bias point V_{GS} = -0.6 Volt and V_{DS} = 3.0 Volt.

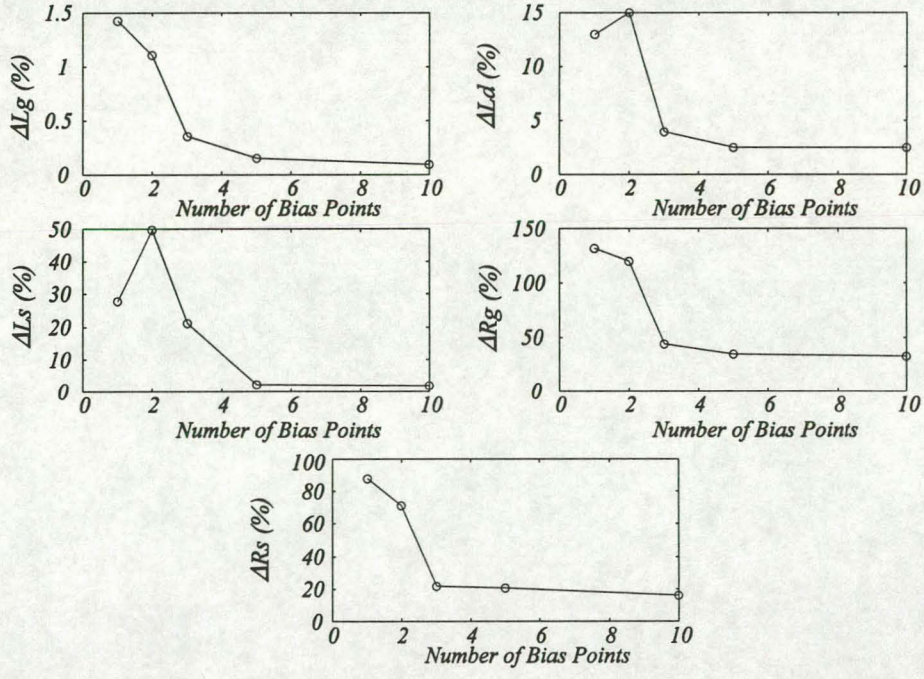


Fig. 3.7 The uncertainty in the MESFET extrinsic elements as a function of the number of bias points used in the extraction.

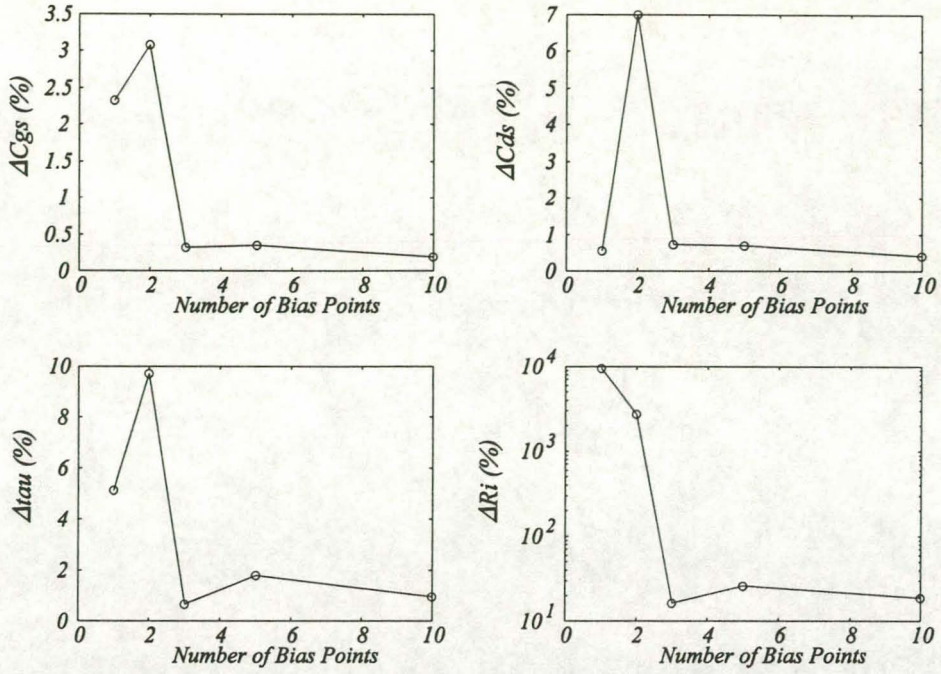


Fig. 3.8 The uncertainty in the pHEMT intrinsic elements as a function of the number of bias points used in the extraction. The intrinsic elements are for the bias point V_{GS} = 0.0 Volt and V_{DS} = 1.5 Volt.

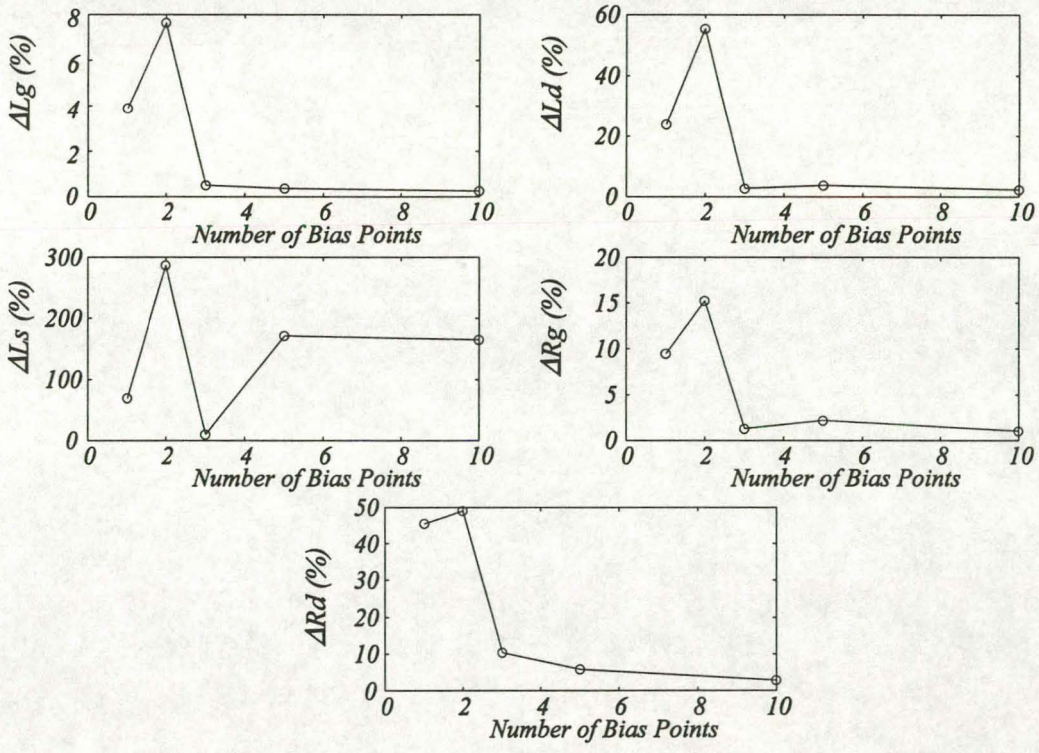


Fig. 3.9 The uncertainty in the pHEMT extrinsic elements as a function of the number of bias points used in the extraction.

A further test for the multi-bias extraction algorithm would be to see if the intrinsic elements extracted from a large number of bias points vary smoothly as a function of the applied bias voltages V_{GS} and V_{DS}. Figures 3.10 to 3.11 show three dimensional plots of the intrinsic elements Cgs, Cgd, τ and Ri for the MESFET and the pHEMT. Compared to the single bias extractions, the multi-bias procedure produces smoother curves for the bias dependent behaviour of the intrinsic model elements, indicating that the procedure is less vulnerable to measurement errors. The variation of τ and Ri with applied bias for the pHEMT is not as smooth as for the MESFET. These two elements have been found by other authors to be extremely difficult to extract [35,38,98] and the pHEMT data is more noisy than that of the MESFET.

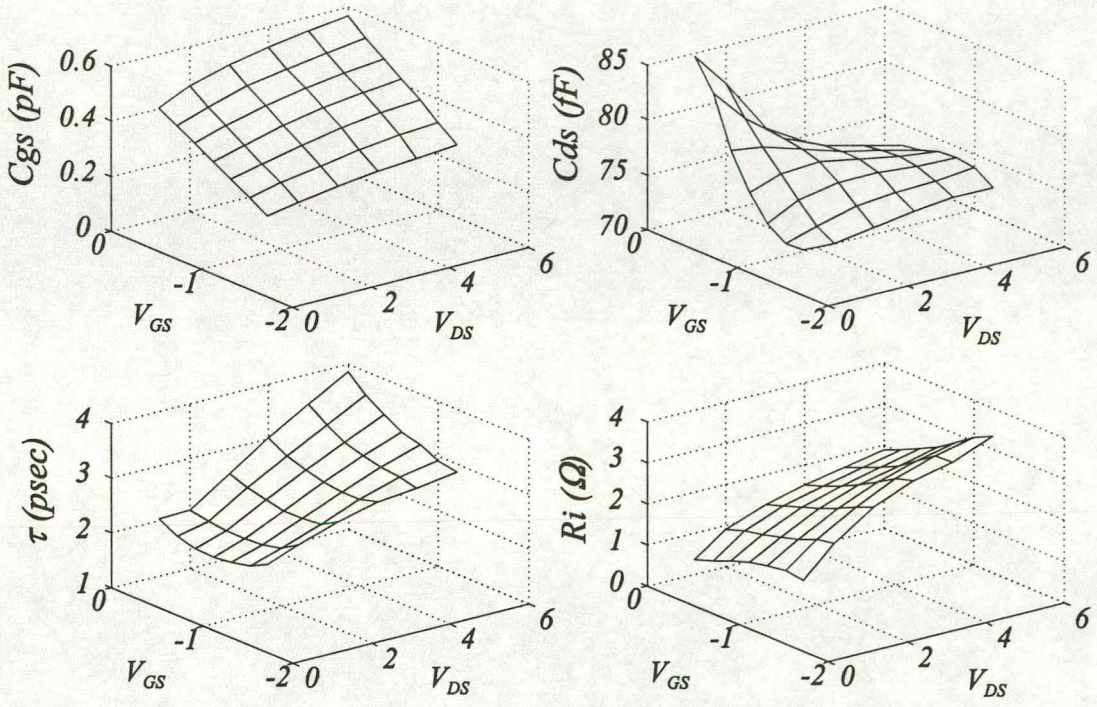


Fig. 3.10 The values of the MESFET intrinsic elements as a function of the bias voltages V_{GS} and V_{DS}.

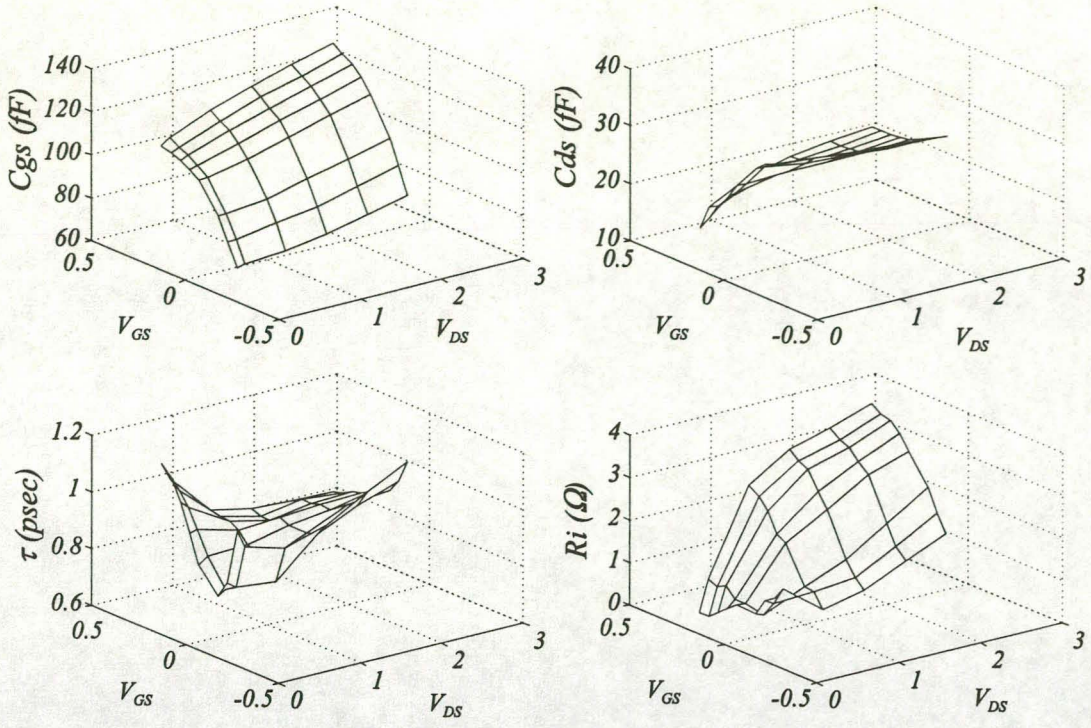


Fig. 3.11 The values of the pHEMT intrinsic elements as a function of the bias voltages V_{GS} and V_{DS}.

A final verification of the multi-bias extraction, is to compare the modelled and measured s-parameters over the whole range of bias points used in the extraction. In order to provide a clear unbiased comparison of the modelling results, the average error in the magnitude and phase of each of the four s-parameters across the frequency band were evaluated for each bias point using equations (3.11) and (3.12). Figures 3.12 to 3.15 show plots of the modelling errors. The errors made in modelling the magnitude and phase of the different s-parameters are small at all the bias points considered in the extraction.

$$e_{magnitude} = \frac{1}{N}\sum_{i=1}^{N} |R_{jk}(\bar{x},\omega_i)| - |S_{jk}(\omega_i)| \tag{3.11}$$

$$e_{phase} = \frac{1}{N}\sum_{i=1}^{N} \angle R_{jk}(\bar{x},\omega_i) - \angle S_{jk}(\omega_i) \tag{3.12}$$

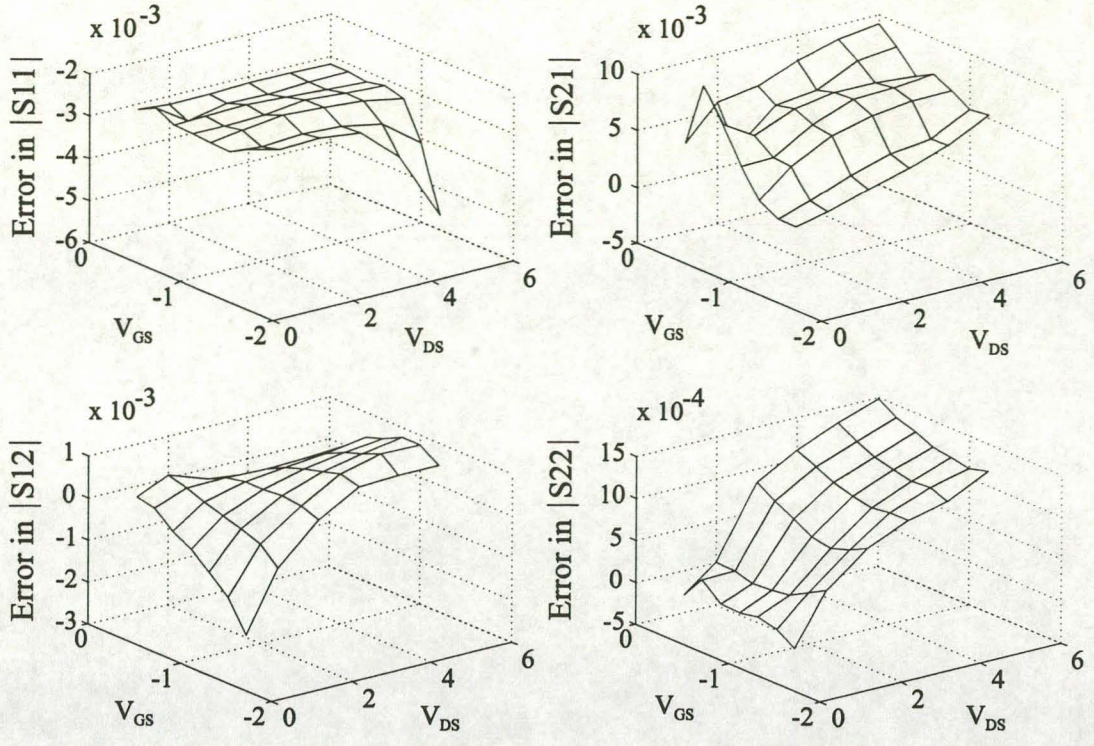


Fig. 3.12 The error made in modelling the magnitude of the MESFET as a function of the bias voltages V_{GS} and V_{DS}.

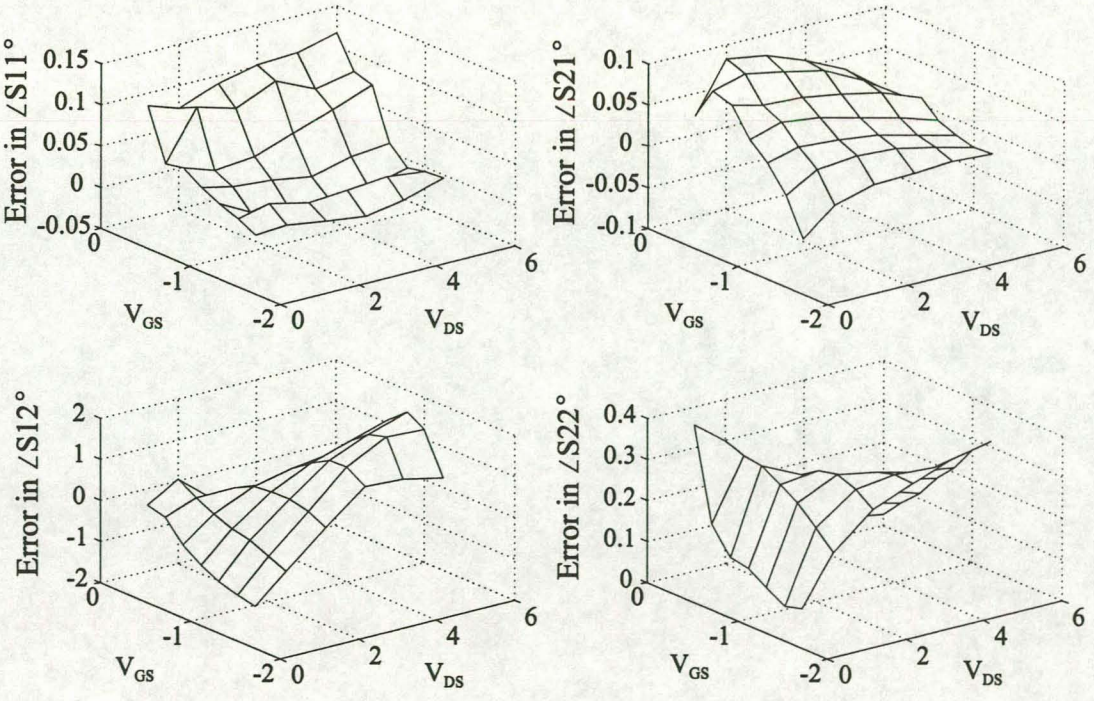


Fig. 3.13 The error made in modelling the phase of the MESFET s-parameters as a function of the bias voltages V_{GS} and V_{DS}. The phase error is expressed in degrees.

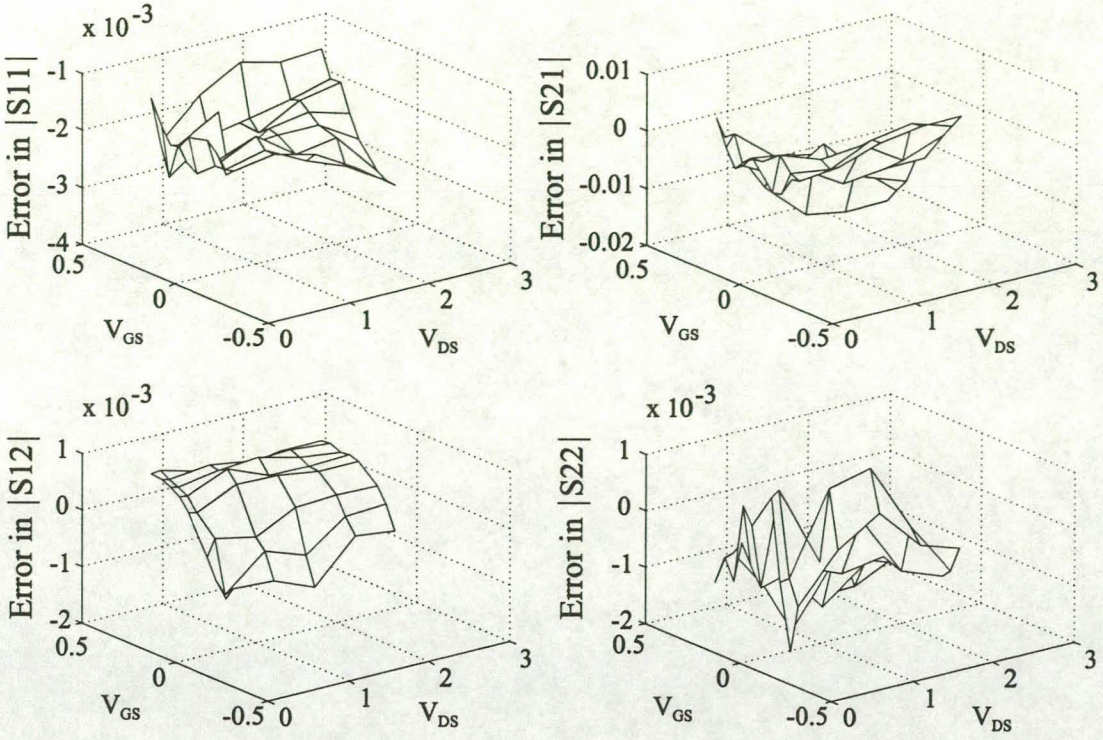


Fig. 3.14 The error made in modelling the magnitude of the pHEMT s-parameters as a function of the bias voltages V_{GS} and V_{DS}.

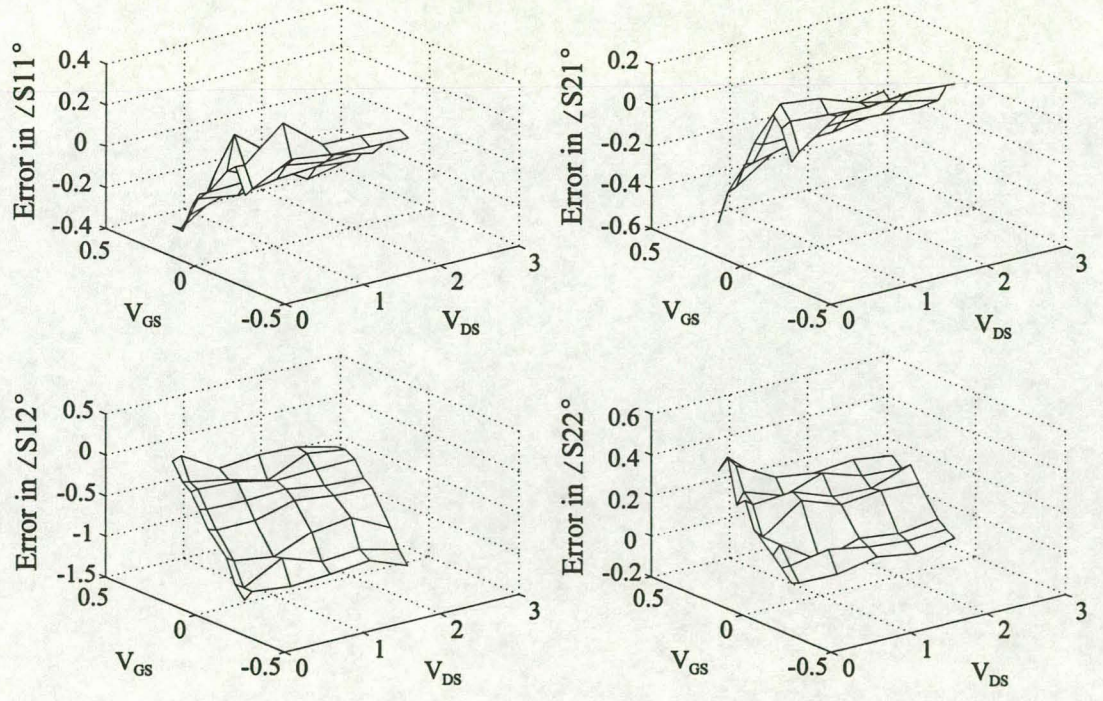


Fig. 3.15 The error made in modelling the phase of the pHEMT s-parameters as a function of the bias voltages V_{GS} and V_{DS}. The errors are expressed in degrees.

3.5 Extending the Multi-Bias Search

In the previous sections it was shown that the multi-bias extraction provides accurate results if a small, but sufficient number of bias points are included in the extraction. The extraction needs to be performed using a wide range of bias points if a complete bias dependent description for a device is to be found. Using the multi-bias extraction, as described in the previous sections, to handle all this data would be time consuming due to the mass of data that must be processed. The following sections will discuss two extensions aimed at making the extraction more efficient. A third section will describe how these extensions can be used to study the bias dependent behaviour of the extrinsic parasitic resistors Rd and Rs.

3.5.1 The Efficient Increase of the Number of Bias Points Used

The direct extraction methods discussed in chapter 1 rely on the fact that if accurate values for the extrinsic elements of a device are known, the intrinsic elements can be calculated using analytical equations. Accurate extrinsic element values can be determined using a relatively small number of bias points in the multi-bias extraction. These elements are de-embedded from the measured data describing other bias points and the corresponding intrinsic elements are calculated. The calculated values serve as an accurate starting point for further optimisations.

Normally the intrinsic elements are determined using the equations derived by Berroth and Bosch [36]. They provide accurate estimates for the intrinsic capacitors Cgs, Cgd, Cds, the transconductance gm and the output resistance Rds. The channel charging resistance Ri and the channel time delay τ are however very badly estimated. A study of the equations of Berroth and Bosch show that the values of these elements will be greatly affected by the errors made in the determination of the other model elements. King used condition numbers to show that these two elements also have a high sensitivity with respect to errors in the measured data [98]. Lin and Kompa [66] presented an alternative set of equations for calculating the intrinsic components. Their equations are derived by using a least squares curve fitting approach. These equations have been found to be far more robust than those of Berroth and Bosch and provide good estimates, even for the insensitive intrinsic elements.

Figures 3.16 to 3.17 show the variation of the intrinsic elements for the MESFET, while figures 3.18 to 3.19 show the same data for the pHEMT. An initial multi-bias search using five bias points were performed, after which the search was expanded using Lin's equations. The extra data was not subjected to any further optimisation. It is clear that the calculations provide very

good estimates of the elements at the other bias points. The calculated element values cannot be distinguished from the elements obtained with the initial optimisation. Figures 3.17 and 3.19 do however indicate that analytical equations have trouble predicting the value of τ at cold biases (V_{DS} = 0.0 Volt). The value of Rds for the MESFET is also predicted incorrectly at one bias point in the linear region. Serious deviations in the calculated results can be corrected in subsequent optimisations. Note that the values of Ri and Cds become negative in the linear region of the I_{DS}-V_{DS} curves. Anholt [35] provides an explanation for the negative Cds values. He suggests that since Cds is only a weak function of bias in the active region of the I_{DS}-V_{DS} plane, it should be fixed at a constant value, even for the linear region of the I_{DS}-V_{DS} plane. His motivation for this is that the modelling errors caused by doing this are relatively small, and it eases the construction of nonlinear models. Since Cds is considered a dominant element in the model, it was found that if its value is not allowed to go negative, it adversely affects the determination of the other model elements. No physical explanation could be found for the negative Ri values, but the smooth and systematic variation of Ri with the applied bias excludes the normal explanation that blames negative element values on measurement uncertainties. If the role of Ri in the model is considered, it is unlikely that the negative values can be given a valid physical explanation. They might however be a side effect of simplifications in the small-signal model in the linear part of the I_{DS}-V_{DS} plane.

The direct calculations provide accurate estimates, but the calculated elements will not yet be at their optimum values and further optimisation is called for. It is argued that if a sufficient number of bias points were used in the initial multi-bias extraction, then the extrinsic elements are already at their optimum values. Further optimisations are therefore confined to the newly calculated intrinsic elements. Figure 3.20 shows the elements Rds, Ri and τ of the MESFET as a function of bias after a decomposition-based optimisation was performed on the intrinsic model elements. The optimiser is able to recover from the errors made in calculating Rds, but τ is unchanged and the variation of Ri with bias is not as smooth as before.

The inability of the optimisation-based and analytical methods to find accurate values for τ at small V_{DS} voltages can be related to the expression used to represent I_{DS} in the small-signal model, namely $V_i \cdot gm \cdot e^{-j\omega\tau}$. The smaller V_{DS} becomes, the smaller the value of gm will be, with gm becoming zero at V_{DS} equal to zero volt. The influence of τ on the measured data will therefore also decrease, till at V_{DS} = 0 Volt, it becomes meaningless.

Figure 3.21 depicts the modelling error (eq. 2.1) as a function of bias. A large increase is noted in the linear region of the device for high V_{GS} voltages. This decrease of modelling accuracy correlates well with the experiences of other researchers [18,35]. Similar results were obtained

for the pHEMT. For the MESFET, there is also a slight increase in the modelling error for large V_{DS} voltages.

The decomposition-based optimiser proved to be robust, but highly inefficient when used to optimise the calculated element values. Rather than subjecting the new data to a full decomposition-based optimiser, a new multi-dimensional second stage optimiser is formulated. This is discussed in the next section.

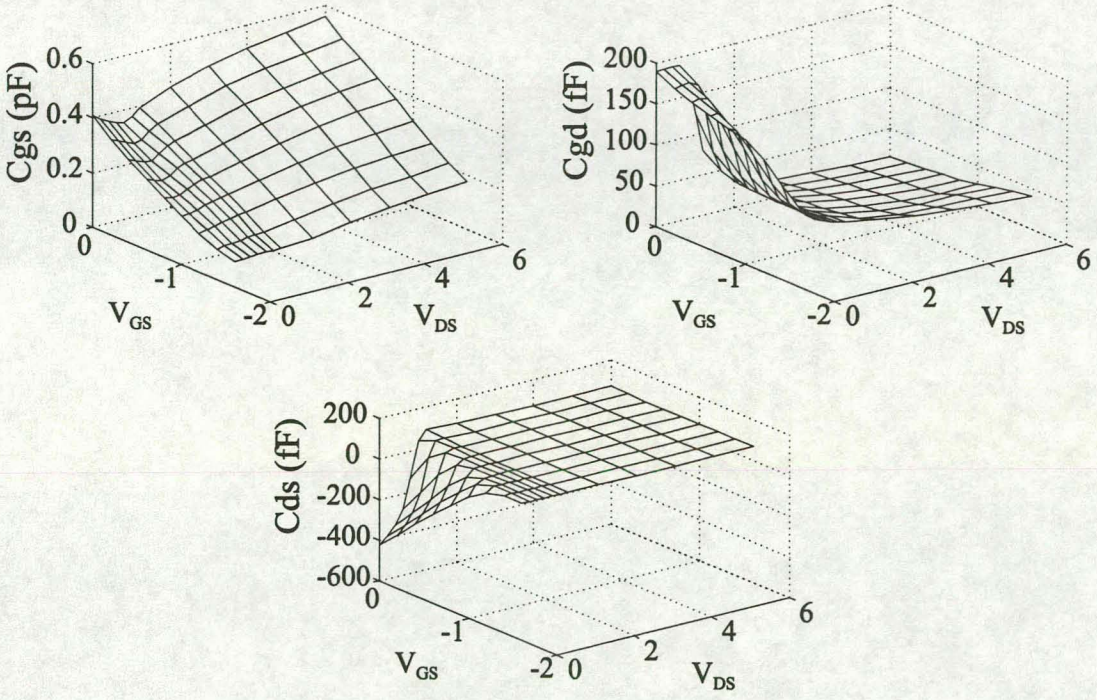


Fig. 3.16 The variation of the MESFET elements Cgs, Cgd and Cds after results of the five bias point multi-bias search were extended with analytical equations.

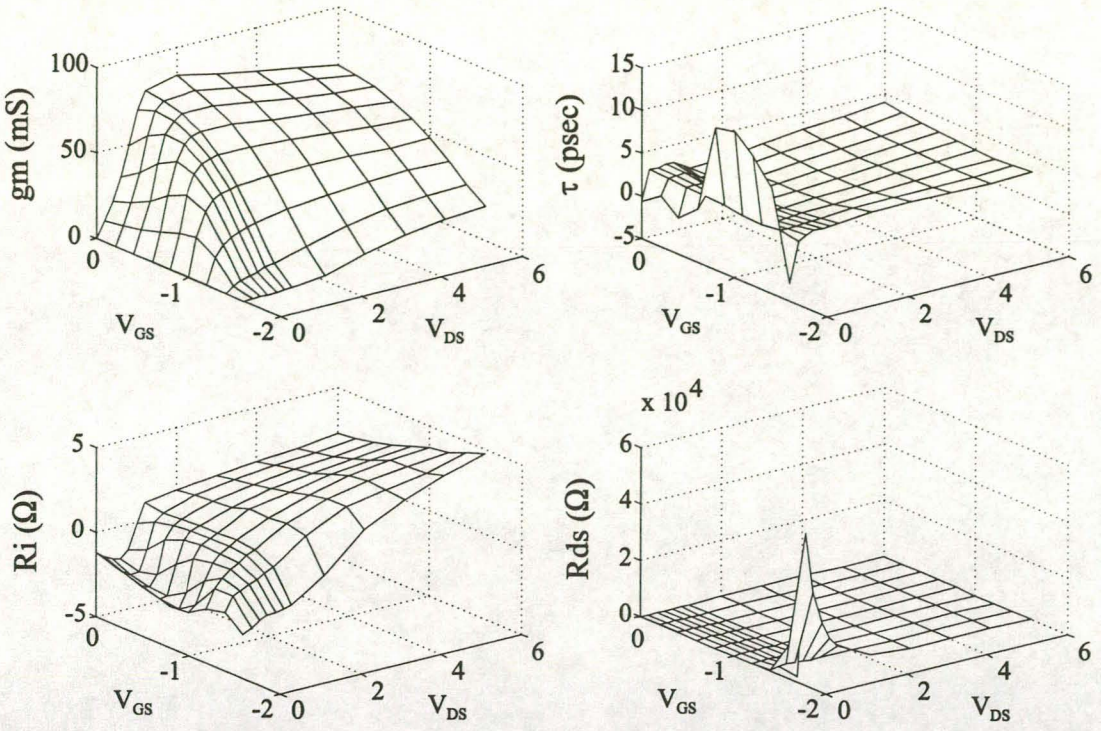


Fig. 3.17 The variation of the MESFET elements gm, τ, Ri and Rds after the results of the five bias point multi-bias search were extended with analytical equations.

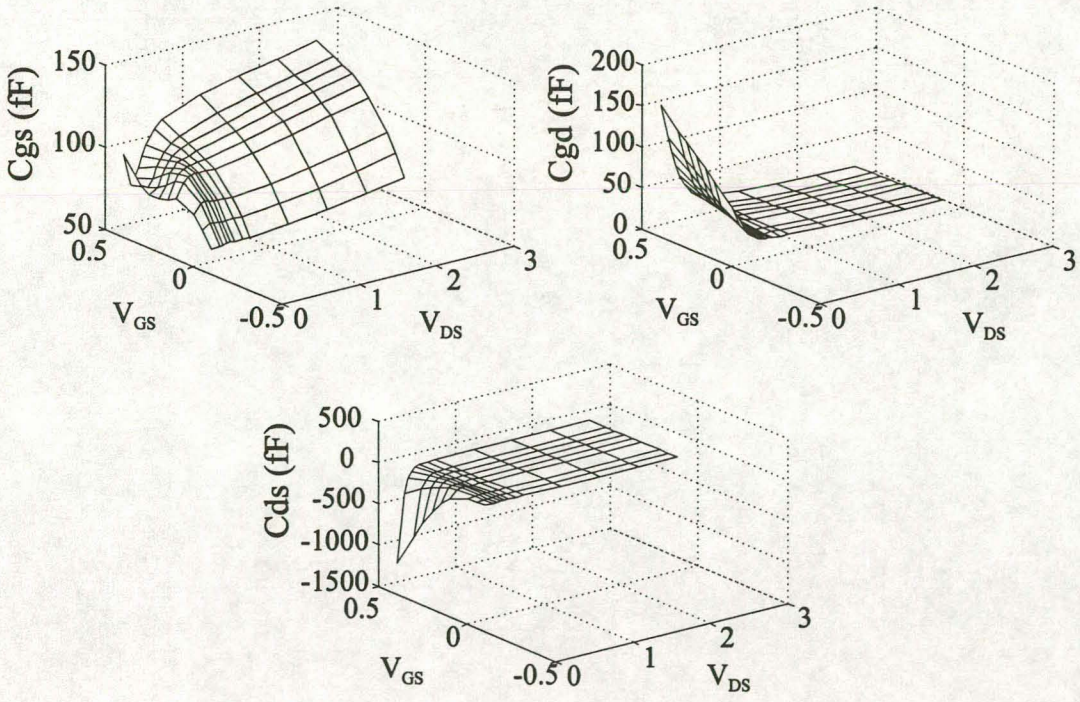


Fig. 3.18 The variation of the pHEMT elements Cgs, Cgd and Cds after results of the five bias point multi-bias search were extended with analytical equations.

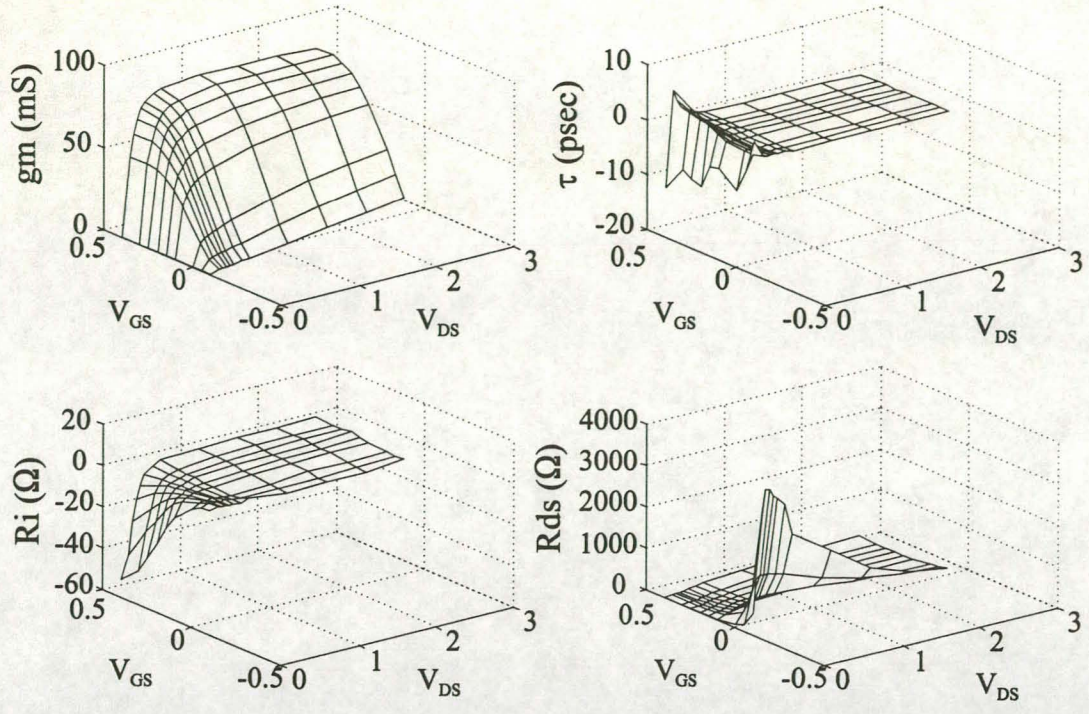


Fig. 3.19 The variation of the pHEMT elements gm, τ, Ri and Rds after results of the five bias point multi-bias search were extended with analytical equations.

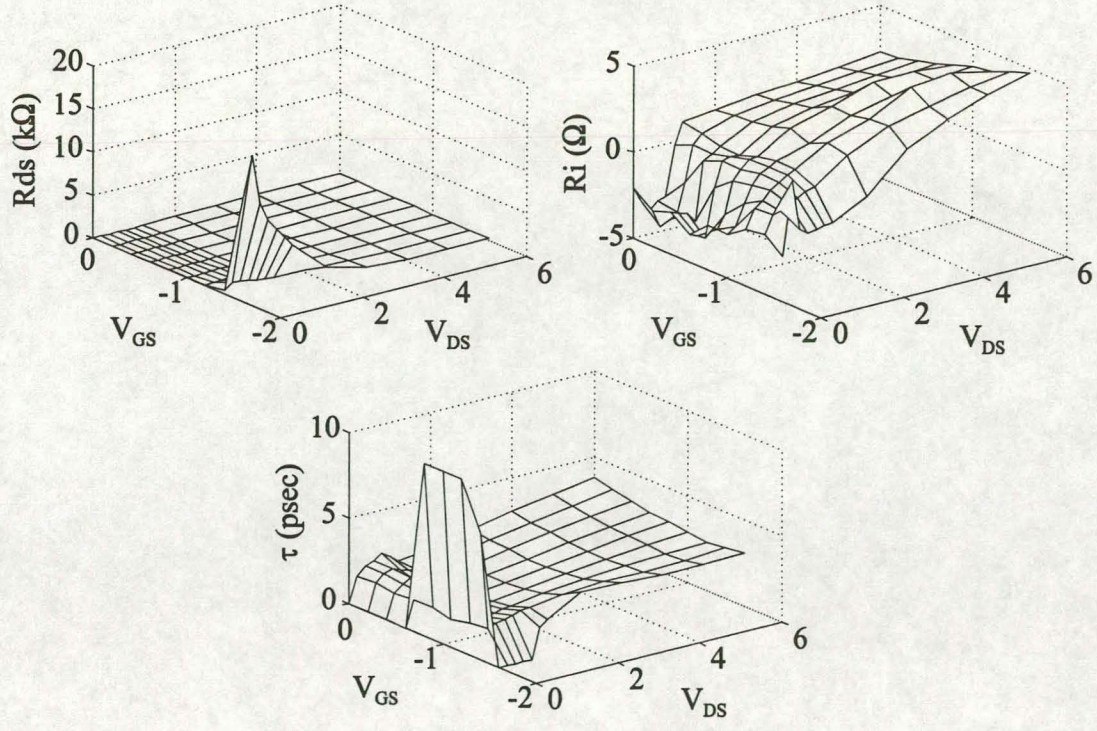


Fig. 3.20 The variation of the MESFET intrinsic elements Rds, Ri and τ as function of bias after being optimised with a decomposition-based algorithm.

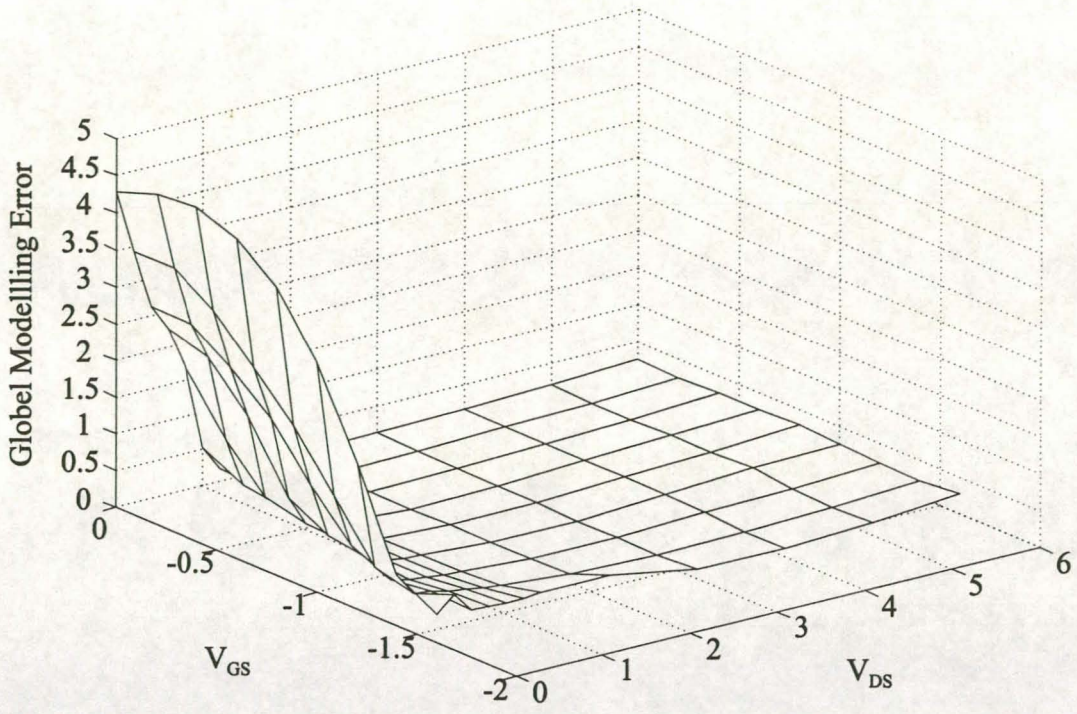


Fig. 3.21 The variation of the modelling error calculated with eq. 2.1 as a function of the applied bias voltages.

3.5.2 The Use of Multi-Dimensional Second Stage Optimisation

The second stage multi-dimensional optimiser serves three different purposes.

- As was shown in the section on convergence, the full decomposition optimiser does not converge to the global minimum, but to a point very close to it. The multi-dimensional optimiser is used to ensure that the global objective function is at a minimum.
- The second use of the multi-dimensional optimiser is to efficiently optimise the calculated intrinsic model elements when the multi-bias search is expanded.
- A third use of the multi-dimensional optimiser is to extract possible bias dependent behaviour in the extrinsic resistors Rd and Rs. This is discussed in the next section.

For the single bias extraction, the second stage optimiser is a conventional multi-variable optimiser, but for the multi-bias problem, the new optimiser can be viewed as a decomposition-based algorithm in which the amount of decomposition has been drastically reduced. The bias independent and bias dependent elements for each bias point, are optimised separately. If there

are M bias points in the extraction, this leads to M+1 sub-optimisation problems. At this low level of decomposition, the optimisation sequence does not matter. The bias independent elements are optimised with respect to an objective function consisting of S_{11}, S_{22} and S_{12} at all the bias points, while the bias dependent elements are optimised with respect to all four the s-parameters that describe their particular bias point. All optimisations are performed using a damped Gauss-Newton algorithm. Three different variations of the multi-dimensional algorithm were implemented. To appreciate the differences between them, a more detailed look at the structure of the decomposition process is called for.

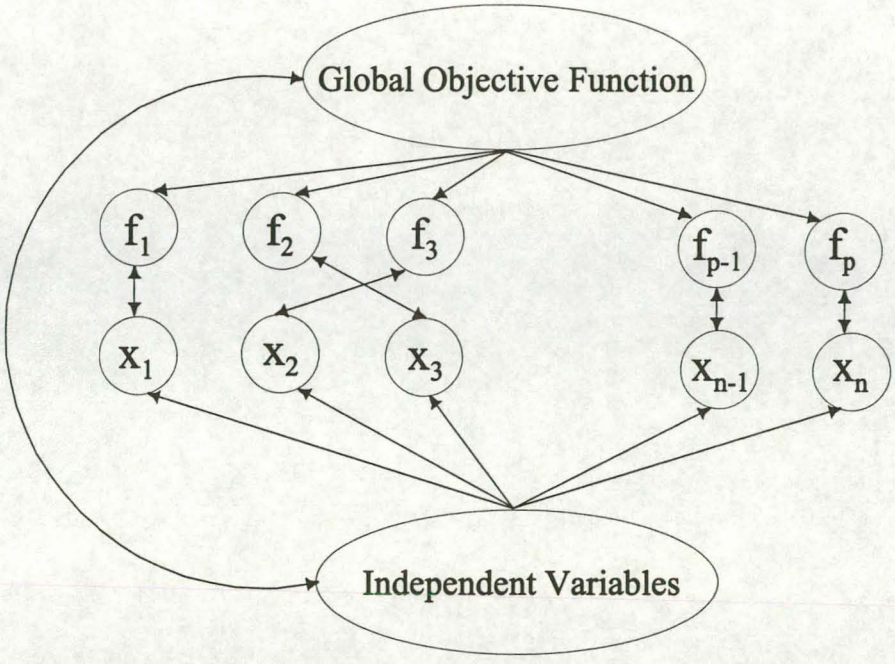


Fig. 3.22 A graphical representation of the decomposition process.

Figure 3.22 shows a graphical representation of the optimisation process. The problem consists of a global objective function and the model elements that are the independent variables of this function. The two inner rows of circles in figure 3.22 are the groups into which the independent variables were divided and the subfunctions into which the global objective function is partitioned. The arrows between them represent the possible subfunction/model element associations. From figure 3.22 it can be seen that the problem is discretized both in terms of its variables and the objective function. By decreasing decomposition, an attempt is made to gradually decrease the amount of discretization until the original optimisation problem consisting of one objective function and its many associated variables are reached again. Bandler [77] proposed that the amount of decomposition is decreased as the minimization

progresses. This certainly makes sense given the convergence behaviour of the algorithm, but the decisions concerning the instances at which decomposition must be lessened introduces an unacceptable number of heuristic choices into the algorithm. It was therefore decided that the transitions from one phase of optimisation to the next will only occur once the current optimisation has converged, and that there will be a direct transition from a high level of decomposition to a low level. A number of strategies using a gradual decrease in decomposition were tried, but none provided more accurate or consistent solutions.

The first implemented multi-dimensional optimiser subjected the variable groups to a conventional damped Gauss-Newton algorithm, making no provision for numerical ill-conditioning. The following equation is used by the Gauss-Newton method to calculate the next set of model elements that will provide a better solution to the problem.

$$\bar{x}^{n+1} = \bar{x}^{n} + \beta\left[2\boldsymbol{J}^{2}\boldsymbol{J}\right]^{-1}\boldsymbol{J}\bar{e} \tag{3.11}$$

β is a scalar damping factor that is adjusted to make sure that the search maintains a downhill direction, n refers to the current iteration of the optimisation, and the rest of the symbols are as defined in the previous sections. Since the Hessian matrix $2\mathbf{J}^{t}\mathbf{J}$ is near singular, the predictions of the less dominant elements in $\bar{x}$ will not be trustworthy and will vary wildly.

The second multi-dimensional optimiser increases the level of decomposition by dividing the model elements into groups that have the same sensitivity. This grouping is performed in such a manner that the Hessian matrix in eq. (3.11) will never have a condition number of more than 10^4 [29]. The objective function is not broken up into subfunctions.

The third variation of the multi-dimensional optimiser makes use of orthogonal transformations to improve the conditioning and speed of convergence [29]. In chapter 2 it is shown that the landscape of the objective function can be visualised as long elliptical contours. The lengths of the principal axis of these contours are proportional to $\sqrt{\lambda_i}$, where λ_i is the i-th eigenvalue of the Hessian matrix. The following orthogonal transformation transforms the elements in $\bar{x}$ to a new set of variables $\bar{y}$.

$$\bar{y} = \boldsymbol{Q}\boldsymbol{D}^{-1/2}\bar{x} \tag{3.12}$$

Q is the matrix whose columns are the eigenvectors of the Hessian matrix and **D** is a diagonal matrix whose elements are the corresponding eigenvalues of the vectors in **Q**. This

transformation provides a set of uncorrelated variables in a landscape that is more circular, and therefore better conditioned. It should be noted that the matrixes **Q** and **D** are both functions of $\bar{x}$ and therefore have to be recomputed as the optimisation progresses. For the FET parameter extraction problem, the added computational burden is not large.

The second and third algorithms were found to be far more sure footed than the conventional Gauss-Newton optimiser, which terminated very quickly if the starting values used where not close to the solution. The second and third multi-dimensional algorithms can reduce the global objective function to low levels, even for inaccurate optimisation starting points. However, neither method converges to the same solution with the consistency of the decomposition-based approach.

Figure 3.23 contains a comparison of the global objective function values returned by the multi-bias decomposition optimiser, and when it is combined with the different second stage multi-dimensional variations of the damped Gauss-Newton algorithm. In each case the decomposition optimiser provides the starting values for the multi-dimensional optimiser. The results in figure 3.23 were obtained using the MESFET data and by repeating the robustness test of section 3.4.

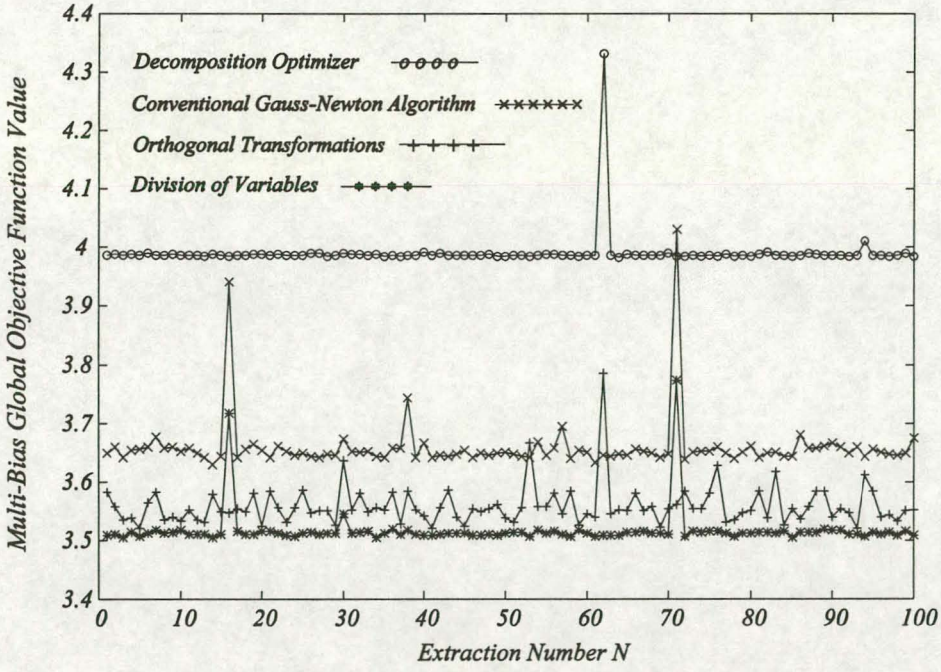


Fig. 3.23 A comparison of the final modelling errors returned by the multi-bias decomposition optimiser and when it is combined with different second stage multi-dimensional variations of the damped Gauss-Newton algorithm.

The error values in figure 3.23 show that the multi-dimensional optimisers only lead to small improvements in the global modelling error. The price that is paid for these improvements can be high. Figure 3.24 shows a histogram of Lg and Ri extracted from the pHEMT data, using first the conventional Gauss-Newton algorithm, and then the variant that divides the variables into groups. Both approaches lead to an increase in the distribution of the extracted elements, especially for the non-dominant elements such as Ri. They also produce multi-modal distributions, even for the dominant model elements. These results also extended to the case where the multi-dimensional optimisers are used to optimise the calculated element values of section 3.5.1. In the saturated region of the device bias plane, the optimised element curves stay smooth, but not in the linear region of the device.

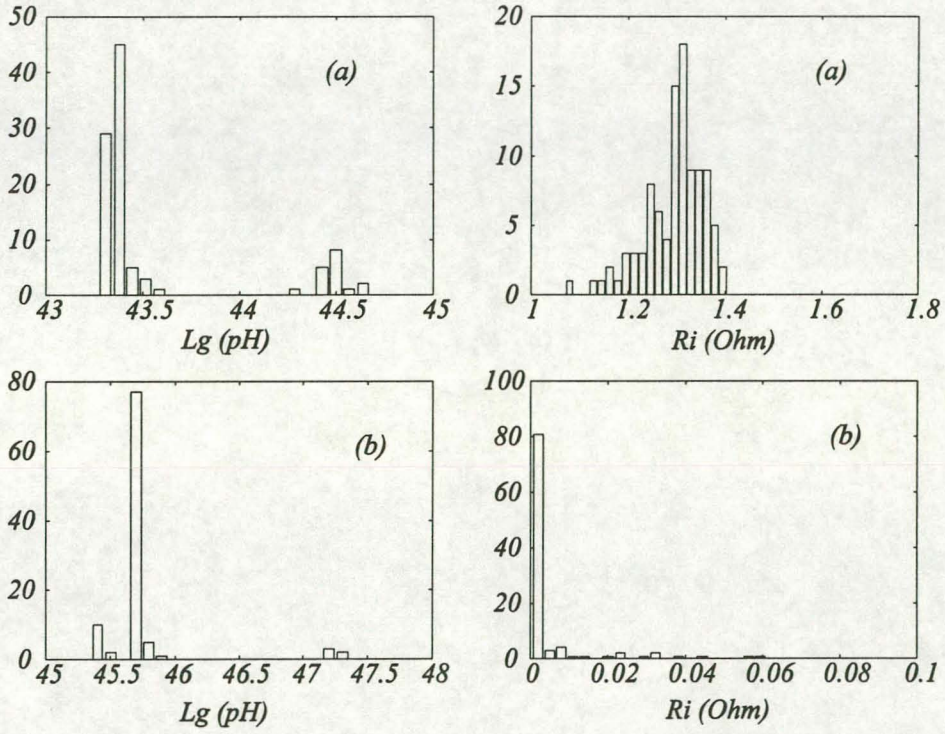


Fig. 3.24 A comparison of the gate inductance and the intrinsic channel resistance Ri after two different second phase multi-dimensional optimisers were used to further optimise a 5 bias point extraction from the pHEMT data. In (a) a conventional Guass-Newton algorithm was used, while in (b) a variation of the Guass-Newton that divides the variables into groups, depending on the sensitivity of the objective function to the variable, was used. Ri describes the bias point V_{GS}=0 Volt and V_{DS}=1.5 Volt.

Table 3.3 A comparison of the average modelling error using only decomposition-based optimisation and when a second phase multi-dimensional optimiser is employed.

GaAs MESFET								
	Bias Point #1 V_{GS} = -0.6 Volt V_{DS} = 3.0 Volt				Bias Point #2 V_{GS} = 0.0 Volt V_{DS} = 6.0 Volt			
	Decomposition Optimiser		Decomposition Optimiser with 2nd Stage Optimiser		Decomposition Optimiser		Decomposition Optimiser with 2nd Stage Optimiser	
	Avg. Amplitude Error	Avg. Phase Error	Avg. Amplitude Error	Avg. Phase Error	Avg. Amplitude Error	Avg. Phase Error	Avg. Amplitude Error	Avg. Phase Error
S11	0.0089	0.3257°	0.0094	0.3124°	0.0100	0.3482°	0.0119	0.3932°
S21	0.0148	0.3367°	0.0140	0.3313°	0.0221	0.5089°	0.0211	0.5063°
S12	0.0009	0.6794°	0.0007	0.5657°	0.0009	1.0996°	0.0008	0.6689°
S22	0.0079	0.4717°	0.0084	0.4250°	0.0215	0.4087°	0.0216	0.3973°
GaAs pHEMT								
	Bias Point #1 V_{GS} = 0.0 Volt V_{DS} = 1.5 Volt				Bias Point #2 V_{GS} = -0.1 Volt V_{DS} = 0.5 Volt			
	Decomposition Optimiser		Decomposition Optimiser with 2nd Stage Optimiser		Decomposition Optimiser		Decomposition Optimiser with 2nd Stage Optimiser	
S11	0.0091	0.5117°	0.0094	0.4970°	0.0070	0.3832°	0.0073	0.4013°
S21	0.0391	0.3820°	0.0404	0.3789°	0.0174	1.2383°	0.0175	1.2476°
S12	0.0031	1.4312°	0.0029	1.3506°	0.0041	1.7250°	0.0040	1.6403°
S22	0.0059	0.4913°	0.0060	0.5021°	0.0080	0.4916°	0.0074	0.4521°

Table 3.3 provides a summary of the modelling errors made in the different s-parameters of the GaAs MESFET and pHEMT with and without the use of a multi-dimensional second phase optimisation. The data was generated using a robustness test with five bias points in the multi-bias extraction. A hundred extractions using random starting values were performed. The comparisons in table 3.3 were generated using the extracted parameter set that corresponds to the largest modelling error. Due to space constraints, only the data for two of the five bias points are listed. Table 3.3 clearly illustrates that the improvements in the modelling accuracy that results from using a second stage multi-dimensional optimiser are small. It is important to note

that the second stage optimiser does not improve the modelling accuracy for all the s-parameters. In several cases there are small decreases in the modelling accuracy. The improvements in the global modelling error shown in figure 3.23 results from improvements in the modelling of only a small number of bias points. This explains why the overall decrease in the global modelling error is so small.

3.5.3 Bias Dependent Behaviour of the Extrinsic Resistors Rd and Rs

Chapter 1 provided an overview of the origin of the parasitic resistances Rd and Rs, the methods used to determine them, and the physical reasons why variations in the values of the parasitic resistances may be observed. In this section, techniques that make use of the robustness of the multi-bias algorithm to extract the bias dependence of Rd and Rs will be discussed.

Ladbrooke [27] defines Rd with the following two equations, depending on whether DC or RF techniques are used to determine it. These two simple equations demonstrate that the two resistances are not necessarily equivalent.

$$R_d = \frac{\partial V}{\partial I_{CH}} \qquad (\textit{microwave resistance}) \tag{3.13}$$

$$R_D = \frac{V}{I_{CH}} \qquad (\textit{dc resistance}) \tag{3.14}$$

V is the voltage across the region between the end of the depletion region and the drain contact and I_{CH} is the current flowing in the channel.

A comparison of different direct RF and DC measurement techniques indicate that they all bias the FET at points where no large fields will exist between the gate and the drain, thus eliminating the formation of Gunn domains. Most direct RF measurements make use of cold FET measurements ($V_{DS} = 0$ Volt) with the gate voltage at a forward bias that is greater than the built-in voltage of the gate Schottky diode ($V_G > V_{bi}$). Reynoso-Hernández *et. al.* [54,55] removed discrepancies between Rd and Rs values found when using DC and RF methods by doing measurements with either the drain or source terminals left open circuited. Once again the fields between the gate and drain are small, preventing the forming of a localized Gunn domain.

Of the two devices tested, only the GaAs MESFET showed possible bias dependent behaviour

in its extrinsic resistors. Apart from the large increase in the modelling error in the linear region of the MESFET (see figure 3.21) there are also increased modelling errors for high V_{DS} voltages. This increase is due to errors made in modelling S_{22} (see figure 3.12) which can be related to the minimum limit that the search imposes on the value of Rd. Rd tends towards a negative value at high V_{DS} voltages.

Two techniques were tested to extract bias dependent values for the extrinsic resistors. In the first attempt, Rd was defined as a bias dependent element in the multi-bias extraction algorithm. This method was not successful due to the effect of the negative resistance values in the circuit model. A negative resistance is a source of energy and this has a large influence on the determination of the other model elements and the convergence of the decomposition-based search. The solid lines in figure 3.25 illustrates the convergence in the values of Rd when it is defined as bias dependent. The extraction was performed using five bias points and a random starting value. The results show the extracted Rd elements varies over an unrealistically large range. The other model elements also take on physically unrealistic values. By making Rd bias dependent, a new group of non-dominant elements are added to the extraction. This, combined with the influence of the negative resistance values, causes the algorithm to exhibit more de-convergence in its third phase of convergence. The algorithms behaviour was further investigated by defining different extrinsic elements as bias independent. None of them caused the algorithm to behave in the same fashion as when Rd is made bias dependent, which leads to the conclusion that it is the negative element values that is the main cause of the unsatisfactory extraction results.

The second approach for determining the bias dependence of Rd starts by using a normal multi-bias decomposition extraction. Once this is completed, the definition of Rd is changed to bias dependent, and the optimisation is continued using either the decomposition algorithm or a multi-dimensional optimisation. When the decomposition-based algorithm is used in the second phase of the optimisation, it continues to minimise the global objective function for about 12 more iterations, after which it starts to de-convergence. The broken line in figure 3.25 shows the convergence behaviour of Rd for this experiment. Note how Rd converges to the same values as found in the first approach, indicating that this is the stable point for the decomposition algorithm.

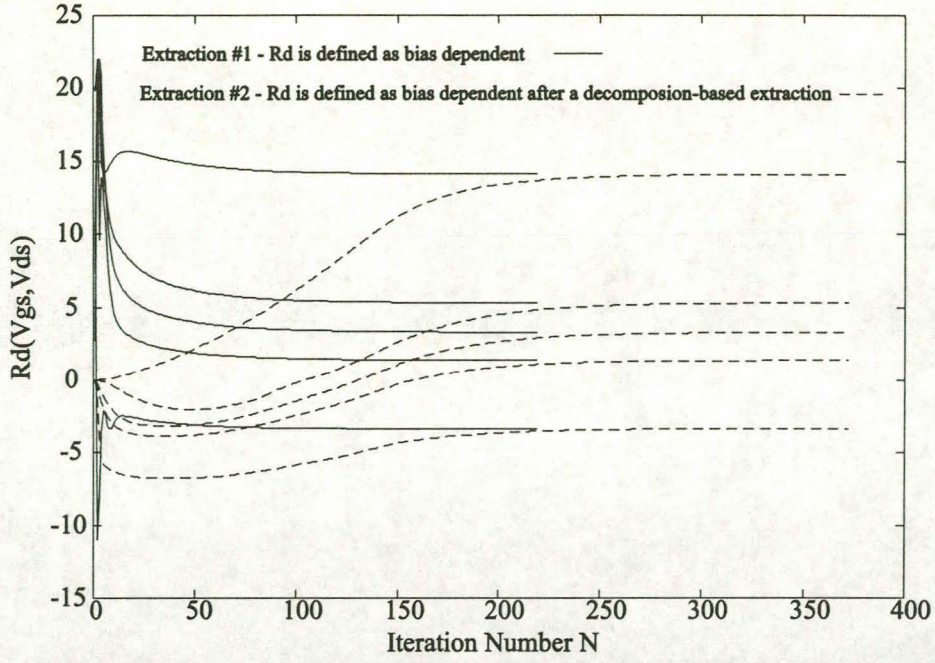


Fig. 3.25 The convergence behaviour of the decomposition-based search for five bias points when Rd is made bias dependent and allowed to assume negative values. In extraction #1 the search is started from a random point and the process converges directly to its stable solution, while in extraction #2 the starting point is solution produced by the normal decomposition-based search when Rd is defined as bias independent. Both approaches converge to the same point.

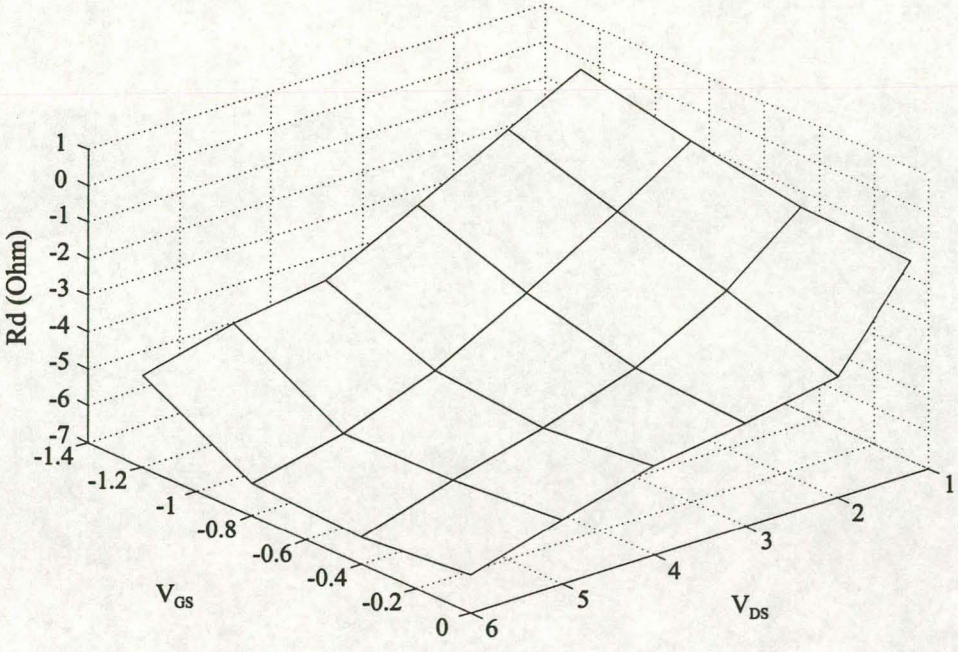


Fig. 3.26 The variation of Rd with the applied bias voltages.

The second phase of the optimisation can also be performed using a multi-dimensional optimiser. This provides more satisfying results because the multi-dimensional algorithms do not break the objective function into subfunctions, thus enforcing the global objective function to be minimised. A test was done using 24 bias points. Figure 3.26 shows the variation of Rd as a function of V_{GS} and V_{DS}. Rd varies smoothly with the applied bias and negative values are found for high V_{DS} voltages. The range of the extracted Rd values is also smaller than previously obtained. The experiment was also performed with both Rs and Rd defined as bias dependent in the second phase search. Rs exhibits very little variation with bias. The maximum change in Rd between the two experiments is 0.4 Ω. Figure 3.27 contains a comparison of the modelling errors before and after the second phase search. By allowing Rd to become bias dependent, the modelling errors at high V_{DS} voltages show improvement.

The negative values of Rd affect the other model elements. Table 3.4 compares the values of the extrinsic elements before and after the multi-dimensional search was used to extract Rd. The changes seen in the element values are not as large as those resulting from defining Rd as bias dependent in the decomposition-based search.

While the origin of the negative resistance values obtained can be explained from physics, the implications of such a resistance in circuit analysis makes it a dangerous element to include in an equivalent circuit model.

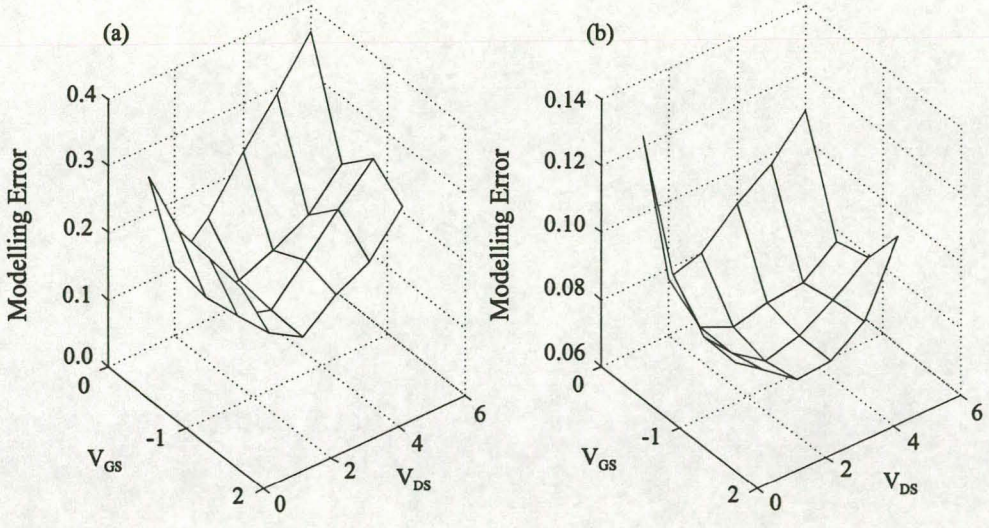


Fig. 3.27 The modelling errors resulting from (a) the conventional decomposition-based extraction with Rd defined as bias independent and (b), when Rd is made bias dependent and extracted using a second phase multi-dimensional optimiser.

Table 3.4 The change in the values of the other extrinsic model element values resulting from second phase multi-dimensional search to extract the bias dependence of Rd.

Extrinsic Element	Before	After	% Change
Lg (pH)	31.171	29.900	4.08 %
Ld (pH)	40.679	47.015	15.58 %
Ls (pH)	4.785	5.030	5.12 %
Rg (Ω)	1.702	2.284	34.20 %
Rs (Ω)	1.179	1.185	0.51 %

3.6 Hybrid Analytical/Decomposition-Based Extractions

It would be unrealistic to expect that one optimisation approach can successfully provide accurate extraction solutions for all instances of device modelling. This section investigates a hybrid approach in which the decomposition-based algorithm is combined with analytical methods. The initial goal of the experiments were to determine if the convergence speed in the flat region of the objective function landscape, surrounding the solution, could be improved.

3.6.1 Creating a Stable Hybrid

Care has to be taken when integrating different optimisation approaches into the decomposition-based algorithm. The power of the decomposition-based optimiser is derived from the interplay of the independent variables during the minimisation of the objective function. If the dynamics of this process are destroyed by the new approach, the hybrid procedure will fail.

A successful hybrid technique can be constructed as follows. When the decomposition-based optimiser has finished a complete cycle of all the sub-optimisation problems, the predicted extrinsic model elements are de-embedded from the measured data, and the analytical equations presented by Lin and Kompa [66] are used to determine the intrinsic model elements. The analytical technique is therefore performed in parallel with the decomposition-based optimiser and produces an alternative set of intrinsic elements. The two processes are linked by accepting the set of intrinsic elements that leads to the lowest global modelling error.

Figure 3.28 illustrates the variation of the global objective function as the search progresses. A multi-bias extraction, making use of the MESFET data and five bias points were performed. In this example, the two parts of the hybrid are run in parallel, but they are not linked. The solid line is the error returned by the decomposition-based optimiser, while the broken line is the objective function value resulting from the analytical calculations. In the early stages of the optimisation, the analytical method provides a faster decimation of the objective function. This changes in the final stages of convergence because the decomposition-based optimiser can adjust the values of the intrinsic model elements individually for the best modelling results. It should be noted that the faster convergence of the analytical section in the early stages of the optimisation is not guaranteed. The reason for this is that the analytical calculations are very much dependent on the values of the intrinsic elements determined by the decomposition-based optimiser.

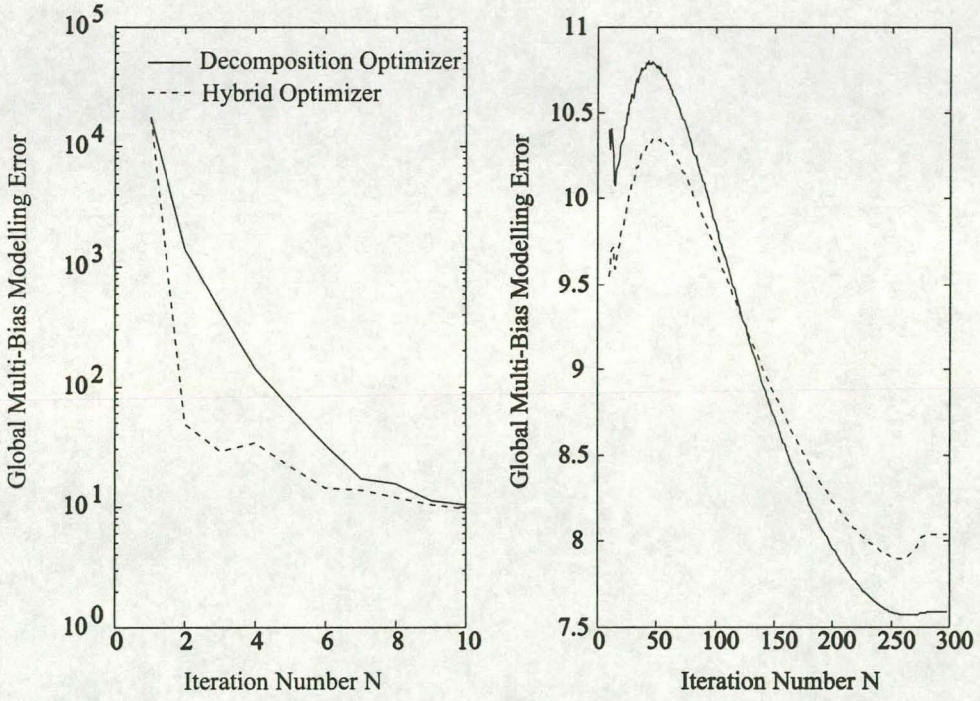


Fig. 3.28 A comparison of the modelling errors returned by the decomposition-based optimiser and the analytical equations when the two are run in parallel, but not coupled to each other. The first graph is the modelling error for the first ten iterations of the search, while the second represent the rest of the search till termination. This example was generated with five bias points in the extraction and the MESFET data.

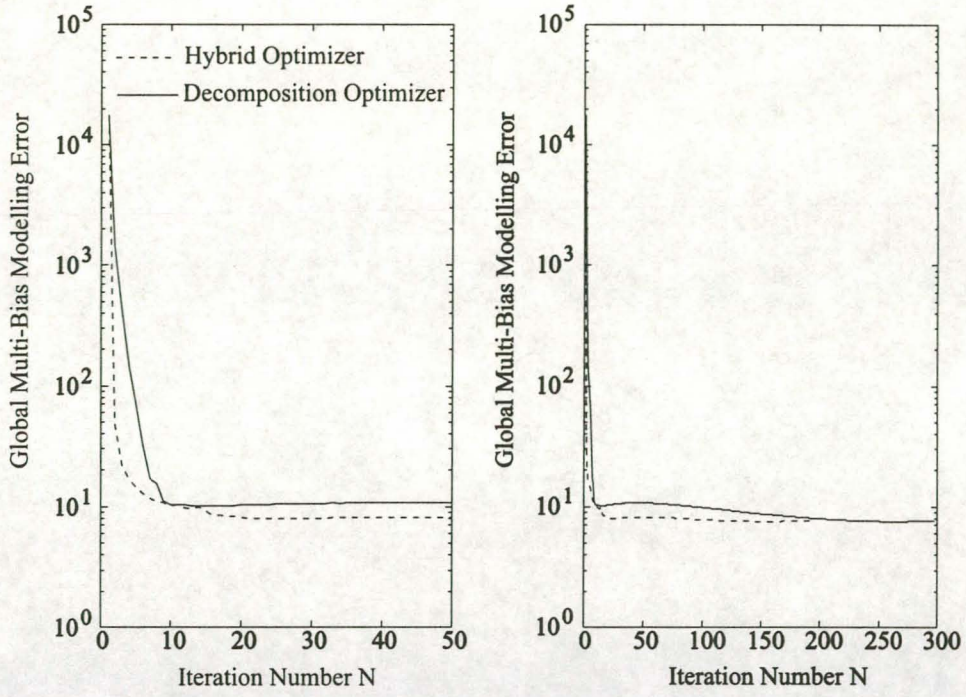


Fig. 3.29 A comparison of the number of iterations needed by the hybrid and the decomposition based optimisers to converge to their final solution. The first part of the graphs zooms in on the first 50 iterations of the search to show the differences between the two algorithms in the initial stages of the extraction.

Figure 3.29 compares the convergence speed of the decomposition-based optimiser and the new hybrid (two sections are now coupled). The extraction consisted of the same multi-bias search that was used for the data in figure 3.28. The hybrid uses fewer iterations to converge to the same solution. The speed improvement is obtained by faster minimisation of the objective function in the early stages of the optimisation process, and because of a stabilising effect that the analytical section has on the method. The decomposition algorithm sometimes performs hill-climbing at a late stage in its convergence. The analytical section of the hybrid dampens this behaviour without destroying the accuracy of the procedure.

The hybrid method provides a simple approach for speeding up the convergence rate of the extraction algorithm. The inclusion of the analytical calculations for determining alternative intrinsic model elements only added a very small computational overhead to the extraction procedure, and do not negatively affect the accuracy of the procedure. It has the added benefit that it increases the robustness of the procedure by adding an extra degree of freedom for dealing with bad optimisation starting values.

3.7 Execution Time and Memory Requirements of the Multi-Bias Optimiser

For the purposes of this thesis, all the examples were executed on a 180 MHz Silicon Graphics workstation with a R10000 processor. The program is written in Fortran 77 and the code was not optimised for speed. The optimisation level of the Fortran compiler used was never set to higher than two. A higher level of compiler optimisation involves the use of approximations in the calculation of mathematical functions.

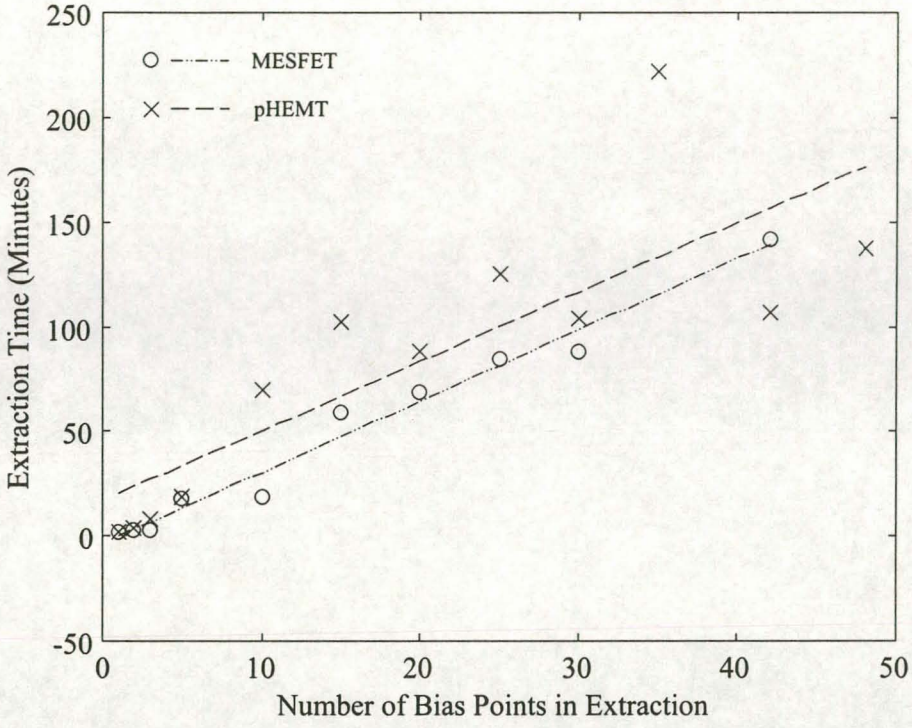


Fig. 3.30 The extraction time for the MESFET and pHEMT as a function of the number of bias points used in the multi-bias search.

The time needed to perform extractions on the MESFET and pHEMT data differed drastically, with the pHEMT proving to be the more difficult extraction problem. The average extraction time for the MESFET, using ten bias points in the search, was 18 minutes 30 seconds. The average extraction time for the pHEMT, using ten bias points in the search, was 1 hour 9 minutes 51 seconds. Figure 3.30 shows the extraction time for the MESFET and pHEMT as a function of the number of bias points used in the extraction. In each case, a linear curve was fitted to the data points using a least squares approach. The MESFET exhibits a very linear relationship between extraction time and the size of the extraction problem. It is suspected that the erratic pHEMT results are due to the simultaneous execution of other numeric codes on the

workstation used. The extraction time was taken as the user time reported by the operating system. While this excludes system time, it is still influenced by the way in which other programs make use of the system resources, as well as the task scheduling performed by the operating system.

The linear relationship between the size of the extraction problem and the number of bias points in the search, can be attributed to the way in which the decomposition approach partitions the problem. The addition of more bias points provide more information for the determination of the extrinsic elements, causing their values to be more uniquely defined. It is expected that this will also increase the complexity of the extrinsic element sub-optimisation problems. However, these functions make up only a small part of the total sub-optimisation problems that need to be solved. This limits their influence on the total extraction time. The determination of bias dependent elements are also kept separate, thus keeping the complexity of the functions that need to be solved constant.

The Fortran 77 program requires 45 MB of memory during execution. Since Fortran 77 does not make provision for dynamic memory management, the program has to be compiled with data structures large enough to accommodate a variety of problems. Higher efficiency in terms of memory management can be obtained by implementing the algorithms in Fortran 90.

3.8 Comparing the Decomposition-Based Optimiser to Other Methods

The previous sections provide extensive evaluations of the robustness and accuracy of the proposed algorithms, but do not explicitly compare them with other extraction procedures that are found in the literature. Using published results, it is possible to draw certain conclusions concerning the speed and accuracy of the different methods. In this section, the decomposition-based algorithm is compared with the systematic multi-dimensional optimisers described by Patterson [29], the procedure of Lin and Kompa [66] and direct extraction methods.

The multi-dimensional optimisers of section 3.5.2 are based on the descriptions of Patterson [29]. Similar results to Patterson, namely that these optimisers are faster and more robust than conventional multi-dimensional optimisers, are obtained. While the multi-dimensional optimisers are able to reduce the objective function to a low value, it does not provide the same tight distribution of element values when tested with random starting values. Both the decomposition-based optimiser and our implementation of the new multi-dimensional optimisers took more iterations to converge than those listed in [29]. The termination conditions used in

[29] are however not known. Since Patterson does not present any histograms of extracted element values, it is not possible to do any further comparisons.

The decomposition-based optimiser is compared to the algorithm proposed by Lin [66] using his published results and the data presented in section 3.6. Lin presents a table of the average and standard deviations of the extracted elements resulting from a robustness test. His single bias extraction exhibits the same behaviour as the single bias extractions performed with the decomposition-based optimiser. The standard deviation of Rg has a magnitude similar to the mean value of Rg, indicating that its absolute value is poorly defined. The value extracted for Rd is also small, similar to the results showed in chapter 2 for Rd values at higher V_{DS} voltages. The extraction accuracy is improved by combining the data from the single bias extraction with data from a cold bias where the gate is held below the pinch-off voltage. However, Lin does not explicitly show whether these results were obtained by integrating the two data sets into one extraction procedure, or if they are the result of two closely coupled single bias extractions. Apart from being accurate, Lin's algorithm is also efficient since the equations used for determining the intrinsic elements can be executed quickly. However, section 3.6 shows that the decomposition-based optimiser can reduce the objective function to a lower value. This is the result of the freedom that the decomposition-based algorithm has to adjust the values of the intrinsic elements independently of each other. The decomposition-based optimiser takes longer to converge than Lin's method, but this is partly due to the differences in the termination criteria used and the flatness of the objective function landscape in the region of the final solution.

Direct parameter extraction algorithms use data from different cold and hot bias points to determine the extrinsic and intrinsic model elements of the small-signal model. They can therefore be viewed as multi-bias algorithms. The procedure described by Tayrani [47] for extracting MESFET or HEMT extrinsic elements was implemented. These elements were then de-embedded from hot s-parameter measurements and the equations of Berroth and Bosch [36] were used to calculate the intrinsic model elements.

Table 3.5 compares the extraction results obtained for the MESFET and pHEMT using the direct extraction with the results from the decomposition-based optimiser. The decomposition-based optimiser results were obtained from the five bias point robustness test described in a previous section. The direct extraction procedure calculates element values for each measurement frequency. Since the elements are frequency independent, there should not be any variation in the element values, but due to the ill-conditioned nature of the problem and small measurement errors, variations do occur. The table lists the average element values, and an extraction uncertainty Δ, which is defined as the difference between the largest and the smallest element

values extracted.

For the calculation of the parasitic inductors and resistors, only measurements above 15 GHz were used, while for the parasitic capacitors, only frequencies below 10 GHz were used. In order to avoid excessive variations in the calculated values of the intrinsic elements, only measurements made at frequencies above 15 GHz were considered. These frequency ranges were chosen after plotting the calculated element values as a function of frequency, and resulted in a considerable reduction of the extraction uncertainty Δ.

Table 3.5 A comparison of the extraction results obtained with a direct extraction method [47] and the new multi-bias decomposition-based optimiser using the MESFET and pHEMT measured data.

	MESFET Bias Point for Intrinsic Elements Vgs = -0.6 Volt Vds = 3.0 Volt				pHEMT Bias Point for Intrinsic Elements Vgs = 0.0 Volt Vds = 1.5 Volt			
	Direct Extraction		Decomposition-Based Optimiser		Direct Extraction		Decomposition-Based Optimiser	
	Avg	Δ	Avg	Δ	Avg	Δ	Avg	Δ
Cgs (fF)	328.672	11.27	354.304	5.8	104.044	8.771	92.158	0.319
Cgd (fF)	32.904	1.607	31.21	0.329	31.698	3.205	33.004	0.028
Cds (fF)	67.997	3.113	74.546	0.54	36.854	8.252	26.287	0.186
gm (mS)	60.52	1.538	65.439	0.001	40.55	1.564	35.214	0.026
τ (psec)	1.965	0.132	2.236	0.01	0.303	0.371	0.850	0.015
Ri (Ω)	4.002	0.888	2.183	0.883	3.876	3.489	1.296	0.333
Rds (Ω)	293.868	53,008	220.793	3.017	410.1	135.98	420.706	0.295
Lg (pH)	30.272	7.242	31.171	0.048	43.274	44.119	43.564	0.156
Ld (pH)	55.044	7.398	40.679	1.007	37.919	11.07	17.326	0.677
Ls (pH)	5.101	0.836	4.785	0.112	6.814	5.537	0.188	0.322
Rg (Ω)	0.757	0.56	1.702	0.59	1.752	1.565	5.743	0.123
Rd (Ω)	1.390	2.061	4.4E-4	0	5.996	3.046	4.484	0.252
Rs (Ω)	0.0257	2.075	1.179	0.239	3.822	3.385	2.0E-4	0
Cpg (fF)	-1.249	——	——	——	15.27	2.662	——	——
Cpd+Cds (fF)	69.381	1.709	——	——	33.917	3.485	——	——

A comparison of the results in table 3.5 indicate that the multi-bias decomposition optimiser consistently produce more unique solutions than the direct extraction algorithm. The variations in the values of the directly determined non-dominant elements have the same magnitude as their average values. For some non-dominant elements, such as Rs in the MESFET, a large number of calculations resulted in non-physical negative element values.

The direct extraction algorithm also produces estimates for the parasitic capacitors. Only the value of Cpg can be determined uniquely since the values of Cds and Cpd cannot be separated. In the case of the MESFET, the Cpg calculations returned negative values, providing a further indication that the parasitic elements are negligible for the MESFET. The 13 element small-signal model thus provides a good approximation for the MESFET. For the pHEMT, the value of Cpg was found to be 15.27 fF, which is not negligible when compared to the value of Cgs. As was previously explained, the maximum measurement frequency was to low for the decomposition-based optimiser to accurately separate the values of these two capacitors.

3.9 Conclusions

The preceding sections provided a detailed discussion about the development of a new multi-bias parameter extraction algorithm. The procedure makes use of the decomposition-based optimiser that was discussed in chapter 2. The multi-bias algorithm combines s-parameters from different bias points into an integrated extraction problem that determines all the model elements simultaneously. By defining the extrinsic elements as bias independent, physical reality is enforced onto the parameter extraction.

The procedure was tested using measured data from the GaAs MESFET and pHEMT devices described in chapter 2. By performing multi-bias extractions using progressively more bias points in each extraction, it is shown that a point is quickly reached where the extracted elements become independent of the measurement deviations in the data. This point is reached if more than five well distributed bias points are used. Robustness tests, performed with a wide range of random optimisation starting values, show that the uncertainty in the values of the extracted elements is small and independent of the starting values used. Extractions using a wide range of bias points produce very smooth curves for the bias dependent descriptions of the intrinsic model elements. The smoothness of these curves indicate that, unlike the single bias extractions, the multi-bias extraction does not follow the measurement imperfections in the data. This is a result of the physical reality that is enforced by defining the extrinsic elements as bias independent.

The decomposition-based multi-bias extraction makes use of the redundant information in the measured data to more accurately determine the values of the non-dominant extrinsic elements. Rg, the parasitic gate resistance, cannot be determined with the single bias extraction, but the multi-bias algorithm extracts it with a high degree of precision.

Extractions at a large number of bias points are needed to obtain accurate bias dependent descriptions of the intrinsic model elements. Due to the mass of data that must be processed, this can be a time consuming procedure. An extension to the multi-bias extraction that makes use of its ability to obtain accurate values for the extrinsic model elements using a small number of bias points was implemented. The extrinsic elements are de-embedded from the data at the other bias points of interest and the corresponding intrinsic elements are calculated using the equations proposed by Lin and Kompa [66]. Very smooth bias dependent curves are once again obtained for the intrinsic elements. The calculated and optimised elements cannot be distinguished from each other.

A comparison of the magnitude and phase of the modelled and measured s-parameters over a wide range of bias points show very small modelling errors. The largest modelling errors are observed in the linear region of the device. These results correlate with those obtained by other authors [18,35].

The multi-bias algorithm shows the same convergence behaviour as the single bias extraction. The algorithm approaches a minimum, after which it moves away (third phase of convergence) to a solution close to the minimum. Experiments were conducted with different multi-dimensional optimisers to determine if the extraction results could be further improved. The multi-dimensional optimisers only lead to small improvements in the modelling quality at the cost of far larger uncertainty in the values of the extracted elements.

The parasitic drain resistance Rd of the MESFET exhibited a large bias dependence, with Rd wanting to become negative at large V_{DS} voltages. Different procedures to extract the bias dependent behaviour of Rd were evaluated. The most successful produced a smooth variation of Rd with applied bias that corresponds with what is expected from the device physics. While it is undeniable that Rd shows a large dependence on the applied bias, it is not clear whether this is because of the type of data used, or a general physical effect. We have no alternative way of verifying the extraction results.

A hybrid approach that combines the decomposition-based algorithm with an analytical method was also investigated. It is shown that this method can increase the algorithm's speed of

convergence without reducing the extraction accuracy. It also has the added benefit of increasing the robustness of the procedure.

The chapter is concluded with a comparison of the proposed multi-bias algorithm with results and other methods that has been presented in the literature. It is found that the multi-bias decomposition-based optimiser is extremely accurate, but that it takes longer for it to converge to its final solution. The method was also compared to results obtained from a direct extraction and consistently produce more unique solutions. The direct extraction results, while not as accurate as the decomposition-based approach, do provide a verification of the extracted model element values.

Chapter 4
Conclusions

4.1 Introduction

This dissertation provided a detailed discussion on the development of a decomposition-based optimiser for extracting multi-bias small signal models from measured s-parameters. Extensive experimental results were presented to illustrate the behaviour and accuracy of the algorithms.

Chapter 2 defined the decomposition-based optimiser and describes how the optimisation problem is divided into subproblems. A key element to the success of the optimiser is the sequence in which these subproblems are solved. It is shown how a principal components sensitivity analysis can be used to calculate an optimum optimisation sequence. This procedure takes the sensitivity of the extraction problem with respect to the different model elements into account. The factors that influence the sensitivity analysis are discussed and patterns in the optimisation sequence are related to the underlying structure of the principal components analysis. Both simulated and measured data were used to evaluate the algorithm. While the procedure proved to be accurate and robust, the extraction results showed the single bias extractions to be influenced by measurement uncertainty. This is evident in the values of the bias independent extrinsic elements that vary with the applied bias and the curves that represent the bias dependent behaviour of the intrinsic elements. These curves are not completely smooth. The chapter is concluded with a discussion on the convergence behaviour of the decomposition-based algorithm.

To overcome the problem of measurement errors, the single bias extraction is expanded in chapter 3 to a multi-bias procedure, that uses data measured at different bias points simultaneously. The extrinsic elements are defined as bias independent. This allows the use of redundant information in the multi-bias data to more accurately define the extraction solution. Experimental results indicate a sharp decrease in the extraction uncertainty when more bias points are included in the extraction process. The devices tested show convergence in both their element values and extraction uncertainty when these are plotted as a function of the number of bias points used in the extraction. The MESFET and pHEMT extraction results indicate that five or more well distributed bias points are needed to obtain results that are independent of measurement imperfections. The curves for the bias dependence of the intrinsic elements are far smoother than for the single bias extractions. The procedure produces models that accurately

represent the s-parameters at all bias points.

When a complete bias dependent description for a device has to be generated, the extraction problems become large due to the mass of data that must be handled. Extensions to the proposed multi-bias algorithm that makes use of analytical methods and multi-dimensional optimisers were discussed and evaluated. The analytical methods greatly improve the speed with which large extractions can be performed, but are prone to errors at certain bias points. The multi-dimensional optimisers provided little improvement in modelling accuracy and lead to increased uncertainty in the extracted element values.

The extensions to the extraction algorithm were also used to investigate the bias dependence of the MESFET parasitic resistors Rs and Rd. The bias dependence of Rd was demonstrated, with Rd taking on negative values at high V_{DS} voltages. While this conforms to predictions made with Monte-Carlo analysis [28], the negative resistance values make the extractions prone to failure and have a large influence on the extraction of the other model elements.

A hybrid decomposition-based optimiser that more closely integrate the analytical methods previously discussed, was also tested. The hybrid optimiser combines the strengths of the two approaches and the results indicate that the hybrid is faster than the decomposition-based optimiser. It also has the added benefit of increasing the robustness of the algorithm by adding an additional degree of freedom for dealing with bad optimisation starting values.

The work is concluded by comparing the decomposition-based algorithm to other published methods and results. Comparisons are made with the systematic optimisation procedures described by Patterson [29] and the multi-plane data fitting technique proposed by Lin and Kompa [66]. The experimental results indicate that the decomposition-based optimiser provides unique solutions that far surpass the capabilities of even advanced multi-dimensional optimisers, such as those described by Patterson [29]. By comparing the test results of the decomposition-based procedure with similar data published by Lin, it is concluded that both procedures are accurate. However, experimental results obtained from the implementation of Lin's analytical equations in the current extraction software, indicate that the decomposition-based optimiser can obtain better modelling accuracies. This is the result of the freedom that the decomposition-based optimiser has to adjust the values of the intrinsic elements independently. It was also found that the optimiser is more resistant to the effects of bad measurement points. This increase in accuracy and robustness comes at the cost of longer extraction times.

Published integrated multi-bias extraction algorithms [58,69,70,73] create one global error

function that contains all the different data sets from which the multi-bias model is to be extracted. These procedures then use conventional optimisation algorithms to minimise this function.

The algorithms that was described in this dissertation show for the first time, that the integrated multi-bias procedure can be efficiently solved using an adaptive decomposition-based optimisation algorithm. The procedure provides highly accurate modelling results. It has a robustness normally only associated with random extraction algorithms, but performs with the speed and efficiency of a gradient search.

4.2 Future Work

The following four areas have been identified for future work on the extraction algorithm.

- Extended hybrid algorithms
- Improvements in the speed of low-level optimisation algorithms
- The inclusion of other small signal-model topologies
- Making provision for the use of other sources of measured data in the multi-bias extraction

The current hybrid algorithm has demonstrated that it can improve the speed of the extraction by providing an alternative estimate for the intrinsic elements. The algorithm can be expanded by also providing analytical equations for estimating the extrinsic elements, given a good estimate of the intrinsic elements. For the 13 element model and single bias data this is a trivial exercise. However, for a successful hybrid, the extrinsic element equations should be as robust as the equations used for the intrinsic elements. They must also be able to handle larger models and multi-bias data.

The speed of the extraction can be further improved by optimising the code used for solving the sub-optimisation problems. These routines are the most frequently called procedures in the programme. Improving their performance will allow larger extraction problems to be handled more effectively.

The current extraction makes use of the conventional small-signal equivalent circuit. In the last few years there have been other suggestions for the intrinsic section of the small-signal model [19]. These new models are aimed at overcoming charge conservation problems encountered

in the construction of nonlinear models. The new models are to be implemented in the software and their extraction results compared with that of the conventional small-signal model.

It would also be logical to expand the multi-bias approach to cater for the inclusion of other types of experimental data, such as DC and noise measurements. This should make it possible to increase the complexity and functionality of the models that are extracted. Other measurements will also help with the accuracy with which the non-dominant elements are determined. This is the most far reaching of the proposed extensions to the current algorithm. It will require a new formulation of the global objective function, which in turn will necessitate changes to the principal components sensitivity analysis.

4.3 Program Implementation

The programs described in the preceding chapters were all implemented in FORTRAN 77. This was done in order to make the software portable, and also because of the availability of compilers for a large variety of platforms. Good, and well tested, mathematical libraries for performing critical functions such as eigenvalue analysis are also available. Future versions of the code will be implemented in FORTRAN 90 or later definitions of this language. This will still allow the code to be portable, while new features in the languages will make the software more efficient.

With increases in the size of the extraction problems that are investigated, the need for pre- and postprocessing programs will become greater. The extraction software is seen a continuously evolving program that will eventually mature to a set of programs that can be used as a tool by a variety of people.

References

[1] D.G. van der Merwe, J.B. Vincent, "GaAs MMICs for commercial microwave products," AP/MTTS-94 Symposium, Stellenbosch, 1994, pp. 51

[2] C.M. Snowden, "Nonlinear Modeling of Power FETs and HBTs," International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering, Vol. 6, No. 4, pp. 219-233

[3] C.M. Snowden, R.R. Pantoja, "GaAs MESFET Physical Models for Process-Orientated Design," IEEE Trans. Microwave Theory and Tech., Vol. 40, No. 7, July 1992, pp. 1401-1409

[4] R. Singh, C.M. Snowden, "Small-Signal Characterisation of Microwave and Millimeter-Wave HEMT's Based on a Physical Model," IEEE Trans. Microwave Theory and Tech., Vol. 44, No. 1, January 1996, pp. 114-121

[5] T. Närhi, "Black-Box Modelling of Nonlinear Devices for Frequency Domain Analysis," 22nd European Microwave Conference, 1992, pp. 1109-1114

[6] F. Filicori, A. Mambrioni, V.A. Monaco, "Large-Signal Narrow Band Quasi-Black-Box Modelling of Microwave Transistors," IEEE MTT-S Digest, 1986, pp. 393-396

[7] F. Filicori, G. Vannini, V.A. Monaco, "A Nonlinear Integral Model of Electron Devices for HB Circuit Analysis," IEEE Trans. Microwave Theory and Tech., Vol. 40, No. 7, July 1992, pp. 1456-1465

[8] H.A. Willing, C. Rauscher, P. de Santis, "A Technique for Predicting Large-Signal Performance of a GaAs MESFET," IEEE Trans. Microwave Theory and Tech., Vol. 26, No. 12, December 1978, pp. 1017-1023

[9] W.R. Curtice, R.L. Camisa, "Self-Consistent GaAs FET Models for Amplifier Design and Device Diagnostics," IEEE Trans. Microwave Theory and Tech., Vol. 32, No. 12, December 1984, pp. 1573-1578

[10] W.R. Curtice, "GaAs MESFET Modeling and Nonlinear CAD," IEEE Trans. Microwave Theory and Tech., Vol. 36, No. 2, February 1988, pp. 220-230

[11] W.R. Curtice, "A MESFET Model for Use in the Design of GaAs Integrated Circuits," IEEE Trans. Microwave Theory and Tech., Vol. 28, No. 5, May 1980, pp. 448-456

[12] J.-M. Dortu, J.-E. Müller, M.Pirola, G. Ghione, "Accurate Large-Signal GaAs MESFET and HEMT Modeling for Power MMIC Amplifier Design," International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering, Vol. 5, No. 3, pp. 195-209

[13] J.J.B. Walker, *High-Power GaAs FET Amplifiers*, Artech House, Norwood, MA, 1993

[14] A. Materka, T. Kacprzak, "Computer Calculation of Large-Signal GaAs FET Amplifier Characteristics," IEEE Trans. Microwave Theory and Tech., Vol. 33, No. 2, 1985, pp. 129-135

[15] H. Statz, P. Newman, I. Smith, R. Pucel, H. Haus, "GaAs FET Device and Circuit Simulation in SPICE," IEEE Trans. Electron Devices, Vol. 34, No. 2, Feb. 1987, pp. 160-169

[16] Y. Tajima, P.D. Miller, "Design of Broad-Band Power GaAs FET Amplifiers," IEEE Trans. Microwave Theory and Tech., Vol. 32, No. 3, March 1984, pp. 261-267

[17] G. Kompa, "Modeling of Dispersive Microwave FET Devices Using a Quasi-Static Approach," International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering, Vol. 5, No. 3, 1995, pp. 173-194

[18] D. Schreurs, *Measurement Based Modelling of Heterojunction Field-Effect Devices for Non-Linear Microwave Circuit Design*, Ph.D. Thesis, Catholic University of Leuven, 1997

[19] Ph. Jansen, D. Schreurs, W. de Raedt, B. Nauwelaers, M. Van Rossum, "Consistent Small-Signal and Large-Signal Extraction Techniques for Heterojunction FET's," IEEE Trans. Microwave Theory and Tech., Vol. 43, No. 1, January 1995, pp. 87-93

[20] A.H. Zaabab, Q.-J. Zhang, M.S. Nakhla, "Device and Circuit-Level Modeling Using Neural Networks with Faster Training Based on Network Sparsity," IEEE Trans. Microwave Theory and Tech., Vol. 45, No. 10, October 1997, pp. 1696-1704

[21] D.E. Root, B. Hughes, "Principles of Nonlinear Active Device Modeling for Circuit Simulation," IEEE MTT ARFTG, December 1988, pp. 3-24

[22] D.E. Root, S. Fan, J. Meyer, "Technology Independent Large Signal Non Quasi-Static FET Models by Direct Construction from Automatically Characterised Device Data," 21st European Microwave Conferernce, 1991, pp. 927-932

[23] I. Angelov, H. Zirath, N. Rorsman, "A New Empirical Nonlinear Model for HEMT and MESFET Devices," IEEE Trans. Microwave Theory and Tech., Vol. 40, No. 12, December 1992, pp. 2258-2266

[24] M.V. Calvo, A.D. Snider, L.P. Dunleavy, "Resolving Capacitor Discrepancies between Large and Small Signal FET Models," IEEE MTT-S Digest, 1995, pp. 1251-1254

[25] S.C. Cripps, "GaAs FET Power Amplifier Design," Matcom Inc., Technical Note 3.2

[26] F. Ali, A. Gupta, *HEMTs & HBTs: Devices, Fabrication, and Circuits*, Norwood, MA, Artech House, 1991

[27] P.H. Ladbrooke, *MMIC Design: GaAs FETs and HEMTs*, Norwood, MA: Artech House,

1989

[28] T. Gonzáles, D. Pardo, "Monte Carlo Determination of the Intrinsic Small-Signal Equivalent Circuit of MESFET's," IEEE Trans. Electron Devices, Vol. 42, No. 4, April 1995, pp. 605-611

[29] A.D. Patterson, V.F. Fusco, J.J. McKeown, J.A.C. Stewart, "A Systematic Optimization Strategy For Microwave Device Modelling," IEEE Trans. Microwave Theory and Tech., Vol. 41, No. 3, March 1993, pp. 395-405

[30] R.L. Vaitkus, "Uncertainty in the Values of GaAs MESFET Equivalent Circuit Elements Extracted from Measured Two-Port Scattering Parameters," Proc. IEEE Cornell Conference on High Speed Semiconductor Devices and Circuits, 1983, pp. 301-308

[31] G. Dambrine, A.Cappy, F. Heliodore, E. Playez, "A New Method for Determining the FET Small-Signal Equivalent Circuit," IEEE Trans. Microwave Theory and Tech., Vol. 36, No. 7, July 1988, pp. 1151-1159

[32] F. Diamand, M. Laviron, "Measurement of Extrinsic Series Elements of a Microwave MESFET Under Zero Current Conditions," 12th European Microwave Conference, Finland, September 1982, pp. 451-456

[33] R. Anholt, S. Swirhum, "Measurement and Analysis of GaAs MESFET Parasitic Capacitances," IEEE Trans. Microwave Theory and Tech., Vol. 38, No. 7, July 1991, pp. 1247-1251

[34] R. Anholt, S. Swirhum, "Equivalent-Circuit Parameter Extraction for Cold GaAs MESFET's," IEEE Trans. Microwave Theory and Tech., Vol. 38, No. 7, July 1991, pp. 1243-1247

[35] R. Anholt, *Electrical and Thermal Characterization of MESFETs, HEMTs, and HBTs*, Artech House, Norwood MA, 1995

[36] M. Berroth, R. Bosch, "Broad-Band Determination of the FET Small-Signal Equivalent Circuit," IEEE Trans. Microwave Theory and Tech., Vol. 38, No. 7, July 1990, pp. 891-895

[37] J.C. Costa, M. Miller, M. Golio, G. Norris, "Fast, Accurate, On-Wafer Extraction of Parasitic Resistances and Inductances in GaAs MESFETs and HEMTs," IEEE MTT-S Digest, 1992, pp. 1011-1014

[38] A. Eskandarian, S. Weinreb, "A Note on Experimental Determination of Small-Signal Equivalent Circuit of Millimeter-Wave FETs," IEEE Trans. Microwave Theory and Tech., Vol. 41, No. 1, January 1993, pp. 159-162

[39] M. Garcia, K. Yhland, H. Zirath, I. Angelov, "Fast, Automatic and Accurate HFET Small-Signal Characterization," Microwave Journal, July 1997, pp. 102-117

[40] B.-S. Kim, S. Nam, K.-S. Seo, "Analytic intrinsic model based parasitic extraction method for HEMTs," Electronics Letters, Vol. 30, No. 12, 9th June 1994, pp. 1005-1006

[41] B-S. Kim, S. Nam, "An iterative parasitic extraction method for HEMT," 25th European Microwave Conference, 1995, pp. 558-561

[42] F. Lenk, R. Doerner, "New Extraction Method for FET Extrinsic Capacitances Using Active Bias Conditions," IEEE MTT-S Digest, 1998, pp. 279-282

[43] C.E. Biber, M.L. Schmatz, T. Morf, U. Lott, E. Morifuji, W. Bächtold, "Technology Independent Degradation of Minimum Noise Figure Due to Pad Parasitics," IEEE MTT-S Digest, 1998, pp. 145-148

[44] G. Leuzzi, K. Deiseroth, F. Giannini, "An Improved Method of Parasitic Extraction for Non-Linear Modelling of Microwave Power MESFET's," 22nd European Microwave Conference, Finland, Vol. 2, 1992, pp. 1206-1210

[45] N. Rorsman, M. Garcia, C. Karlsson, H. Zirath, "Accurate Small-Signal Modeling of HFET's for Millimeter-Wave Applications," IEEE Trans. Microwave Theory and Tech., Vol. 44, No. 3, March 1996, pp. 432-437

[46] W. Stiebler, M. Matthes, G. Böck, T. Koppel, A. Schäfer, "Bias-Dependant "Cold-(H)FET" Modelling," IEEE MTT.-S Digest, 1996, pp. 1313-1316

[47] R. Tayrani, J.E. Gerber, T. Daniel, R.S. Pengelly, U.L. Rohde, "A new and reliable direct parasitic extraction method for MESFETs and HEMTs," 23rd European Microwave Conference, 1993, pp. 451-453

[48] H.O. Vickes, "Determination of Intrinsic FET Parameters Using Circuit Partitioning Approach," IEEE Trans. Microwave Theory and Tech., Vol. 39, No. 2, February 1991, pp. 363-366

[49] P.M. White, R.M. Healy, "Improved Equivalent Circuit for Determination of MESFET and HEMT Parasitic Capacitances from "Coldfet" Measurements," IEEE Microwave and Guided Wave Letters, Vol. 3, No. 12, December 1993, pp. 453-454

[50] F. Blanco, "New Techniques to Measure Parasitic Resistances in Schottky-Gate FETs," Solid-State Electronics, Vol. 37, No. 3, 1994, pp. 451-458

[51] V.I. Cojocaru, T.J. Brazil, "Parasitic Resistance Extraction Errors with Implications for FET Model Accuracy Around Vds = 0," IEEE MTT-S Digest, 1997, pp. 1599-1602

[52] P. Debi, L. Martens, "Fast and Accurate Extraction of Parasitic Resistances for Nonlinear GaAs MESFET Device Models," IEEE Trans. Microwave Theory and Tech., Vol. 42, No. 12, December 1995, pp. 2239-2242

[53] K.W. Eccleston, "Method for determination of parasitic resistances in microwave power MESFETs," IEE Proceedings-G, Vol. 136, No. 6, December 1989, pp. 358-360

[54] J.A. Reynoso-Hernàndez, F.E. Rangel-Patiño, "DC and RF Techniques for Computing Access Resistances in Microwave FET's," IEEE MTT-S Digest, 1996, pp. 1711-1714

[55] J.A. Reynoso-Hernàndez, F.E. Rangel-Patiño, J. Perdomo, "Full RF Characterization for Extracting the Small-Signal Equivalent Circuit in Microwave FET's," IEEE Trans. Microwave Theory and Tech., Vol. 44, No. 12, December 1996, pp. 2625-2633

[56] V. Sommer, "A New Method to Determine the Source Resistance of FET from Measured S-Parameters Under Active-Bias Conditions," IEEE Trans. Microwave Theory and Tech., Vol. 43, No. 3, March 1995, pp. 504-510

[57] R.Vogel, "Determination of the MESFET Resistive Parameters Using RF-Wafer Probing," 17th European Microwave Conference, 1987, pp. 616-621

[58] J.W. Bandler, S.H. Chen, S. Daijavad, "Microwave Device Modelling Using Efficient l_1 Optimization: A Novel Approach," IEEE Trans. Microwave Theory and Tech., Vol. 34, No. 12, December 1986, pp. 1282-1292

[59] J.P. Bridge, P.H. Ladbrooke, A.J. Hill, "Characterisation of GaAs FET and HEMT chips and packages for accurate hybrid circuit design," IEE Proceedings-H, Vol. 139, No. 4, August 1992, pp. 330-336

[60] V.F. Fusco, "Symbolic transfer functions for MESFET small-signal parameter extraction," IEE Proceedings-G, Vol. 138, No. 2, April 1991, pp. 217-221

[61] A.K. Jastrzebski, A. Davies, "Extraction of Linear and Non-Linear MESFET Models," 21st European Microwave Conference, 1991, pp. 945-952

[62] G. Kompa, F. Lin, "FET Modelling using an Analytic Extraction Method Based on Broadband S-Parameter Measurements," 20th European Microwave Conference, pp. 778-783

[63] G. Kompa, M. Novotny, "Highly Consistent FET Model Parameter Extraction Based on Broadband S-Parameter Measurements, "IEEE MTT-S Digest, 1992, pp. 293-296

[64] H. Kondoh, "An Accurate FET Modelling from Measured S-Parameters," IEEE MTT-S Digest, 1986, pp. 377-380

[65] M.S. Leong, P.S. Kooi, T.S. Yeo, B.L. Ooi, "A Hybrid Gauss-Newton-Simulated Annealing Optimizer for Extraction of MESFET Equivalent Circuit Elements," Microwave and Optical Technology Letters, Vol. 6, No. 8, June 1993, pp. 461-466

[66] F. Lin, G. Kompa, "FET Model Parameter Extraction Based on Optimization With Multiplane Data-Fitting and Bidirectional Search - A New Concept," IEEE Trans. Microwave Theory and Tech., Vol. 42, No. 7, July 1994, pp. 1114-1121

[67] K. Shirakawa, H. Oikawa, T. Shimura, Y. Kawasaki, Y. Ohashi, T. Saito, Y. Daido, "An

Approach to Determining an Equivalent Circuit for HEMT's," IEEE Trans. Microwave Theory and Tech., Vol. 43, No. 3, March 1995, pp. 499-503

[68] C. van Niekerk, P. Meyer, "A New Approach for the Extraction of an FET Equivalent Circuit from Measured S-Parameters," Microwave and Optical Technology Letters, Vol. 11, No. 5, April 1996, pp. 281-284

[69] Q. Cai, J. Gerber, U.L. Rohde, T. Daniel, "HBT High-Frequency Modeling and Integrated Parameter Extraction," IEEE Trans. Microwave Theory and Tech., Vol. 45, No. 12, December 1997, pp. 2493-2501

[70] S. Lee, "Fast and Efficient Extraction of HBT Model Parameters Using Multibias S-Parameter Sets," IEEE Trans. Microwave Theory and Tech., Vol. 44, No. 8, August 1996, pp. 1499-1502

[71] G.L. Bilbro, M.B. Steer, R.J. Trew, C.R. Chang, S.G. Skaggs, "Extraction of the Parameters of Equivalent Circuits of Microwave Transistors Using Tree Annealing," IEEE Trans. Microwave Theory and Tech., Vol. 38, November 1990, pp. 1711-1718

[72] R. Menozzi, A. Piazzi, F. Contini, "Small-Signal Modeling for Microwave FET Linear Circuits Based on a Genetic Algorithm," IEEE Trans. on Circuits and Systems - I:Fundamental Theory and Applications, Vol. 43, No. 10, October 1996, pp. 839-847

[73] A. Ghazinour, R.H. Jansen, "Robust, Model-Independant Generation of Intrinsic Characteristics and Multi-bias Parameter Extraction for MESFETS/HEMTS," IEEE MTT-S Digest, 1998, pp. 149-152

[74] I.O. Bohachevsky. M.E. Johnson, M.L. Stein, "Generalized Simulated Annealing for Function Optimization," Technometrics, Vol. 28, No. 3, August 1986, pp. 209-216

[75] K.S. Tang, K.F. Man, S. Kwong, Q. He, "Genetic Algorithms and their Applications," IEEE Signal Processing Magazine, November 1996, pp. 22-37

[76] P.J.M. Laarhoven, E.H.L. Aarts, *Simulated Annealing Theory and Applications*, D. Reidel, Dordrecht, 1988

[77] J.W. Bandler, Q.-J. Zhang, "An Automatic Decomposition Approach to Optimization of Large Microwave Systems," IEEE Trans. Microwave Theory and Tech., Vol. 35, No. 12, December 1987, pp. 1231-1239

[78] C. van Niekerk, P, Meyer, D. Schreurs, P. Winson, "A Robust Integrated Multi-Bias Parameter Extraction Method for MESFET and HEMT Models," Submitted for Publication to the IEEE Transactions on Microwave Theory and Tech.

[79] Van Niekerk, P. Meyer, "Performance and Limitations of Decomposition-Based Parameter-Extraction Procedures for FET Small-Signal Models," IEEE Trans. Microwave Theory and Tech., Vol 46, No. 11, November 1998, pp. 1620-1627

[80] S.W. Director, *Decomposition of Large-Scale Problems*, Amsterdam: North-Holland, 1973, pp. 94-117

[81] J.J. McKeown, "On Algorithms for Sums of Squares Problems," Towards Global Optimization, Edited by L.C.W. Dixon and G.P. Szegö, North-Holland/American Elsevier, 1974

[82] K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space," Edinburgh and Dublin Philosophical Magazine and Journal of Science, Vol. 2, Part SER, 1901, pp. 559-572

[83] D.F. Morrison, *Multivariate Statistical Methods*, New York, McGraw-Hill, 1967

[84] C.-H. Chen, *Statistical Pattern Recognition*, Rochelle Park, New Jersey, Hayden, 1973

[85] E. Kreyszig, *Advanced Engineering Mathematics*, 6th Edition, John Wiley & Sons, New York, 1988

[86] J.J. McKeown, A. Nag, "An Application of Optimization Techniques to the Design of an Optical Filter," Optimization in Action, Edited by Dixon, Academic Press, London, 1976

[87] T.D. Wickens, *The Geometry of Multivariate Statistics*, Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1995

[88] A. Jennings, J.J. McKeown, *Matrix Computation*, 2nd Edition, John Wiley & Sons, Chichester, England, 1992

[89] J.J. McKeown, D. Meegan, D. Sprevak, *An Introduction to Unconstrained Optimisation*, Adam Hilger, Bristol, 1990

[90] B.T. Smith, J.M. Boyle, J.J. Dongarra, B.S. Garbow, Y. Ikebe, V.C. Klema, C.B. Moler, *Matrix Eigensystem Routines – EISPACK Guide*, Lecture Notes in Computer Science, Vol. 6, 2nd Edition, Springer-Verlag, 1976

[91] Fujitsu Semiconductors Data Book, 1993

[92] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd Edition, McGraw-Hill, New York, 1991

[93] P.B. Winson, Personal Communication

[94] Y. Baeyens, T. Skrabka, M. van Hove, W De Raedt, B. Nauwelaers, M. Van Rossum,"Performance of 0.2 μm planar doped pseudomorphic and lattice matched HEMTs on GaAs and InP, " Proc. 23rd European Solid State Device Research Conference (ESSDERC '93), pp. 753-756, 1993

[95] D. Schreurs, Personal Communication

[96] D. Schreurs, Y. Baeyens, B.K.J.C. Nauwelaers, W. De Raedt, M. Van Hove, M. Van Rossum, "S-Parameter Measurement Based Quasistatic Large-Signal Cold HEMT Model for Resistive Mixer Design," International Journal of Microwave and Millimeter-Wave Computer-Aided Engineering, Vol. 6, No. 4, 1996, pp. 250-258

[97] M. Novotny, G. Kompa, "Unique and physically meaningful extraction of the bias-dependent series resistors of a 0.15 μm PHEMT demands extremely broadband and highly accurate measurements," IEEE MTT-S Digest, 1996, pp. 1715-1718

[98] F.D. King, P. Winson, A.D. Snider, L. Dunleavy, D.P. Levinson, "Math Methods in Transistor Modeling: Condition Numbers for Parameter Extraction," IEEE Trans. Microwave Theory and Tech., Vol. 46, No. 9, September 1998, pp. 1313-1314